evidence development

Systematic Review

Clinical and Economic Systematic Reviews in Treatment Naïve Advanced/Metastatic Renal Cell Carcinoma

Version 3.0

HTA Style Report

December 2010

Confidential: For Internal Use Only

Conducted by: Heron Evidence Development Ltd

Table of Contents

List of Abbreviations

Abbreviation	Full definition
AE	Adverse event
ALT	Alanine transaminase
ASCO	American Society of Clinical Oncology
ASCO-GU	ASCO-Genitourinary
AST	Aspartate transaminase
AVOREN	Avastin for Renal Cell Cancer trial
AWMSG	All Wales Medicines Strategy Group
BAY43-9006	Sorafenib
Bev	Bevacizumab
BL	Baseline
BSC	Best supportive care
CADTH	The Canadian Agency for Drugs and Technologies in Health
CALGB	Cancer and Leukemia Group B trial
CCI-779	Temsirolimus
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
CHF	Congestive heart failure
CI	Confidence interval
CONSORT	Consolidated Standards of Reporting Trials
CR	Complete response
CRECY	The Cancer du Rein et Cytokines trial
CSR	Clinical study report
DCR	Disease control rate
DFS	Disease free survival
DOR	Duration of response
DVT	Deep venous thrombosis
ECCO	European Conference for Clinical Oncology
ECOG	Eastern Cooperative Oncology Group
EMBASE	Excerpta Medica Database
EMEA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
EORTC QLQ C-30	EORTC Quality of Life-Questionnaire-C30
EQ-5D	EuroQoL Questionnaire
EQ-VAS	EuroQoL -Visual Analog Scale
ESMO	European Society for Medical Oncology
FACT–G	Functional assessment of cancer therapy – general scale
FKSI	FACT – kidney symptom index
FKSI-15 Index	FACT-Kidney Symptom Index–15 item scale
FKSI-DRS Index	FACT-Kidney Symptom Index–Disease-related Symptom subscale
GSK	GlaxoSmithKline
GW786034	Pazopanib
HR	Hazard ratio
HRQoL	Health Related Quality of Life
IFN	Interferon alpha
IL-2	Interleukin 2
ISRCTN	International Standard Randomised Controlled Trial Number
ITT	Intention to treat analysis
KPS	Karnofsky Performance Status
LSMs	Least square means
MEDLINE	Medical Literature Analysis and Retrieval System Online
MPA	Medroxyprogesterone
MRC RE01	Medical Research Council RE01 trial
MSKCC	Memorial Sloan-Kettering Cancer Centre
MTA	Multiple technology appraisal
mTOR	Mammalian target of rapamycin
N/A	Not available
NCC-RCC	Non-clear cell renal carcinoma
NE	Number of evaluable patients
NICE	National Institute for Health and Clinical Excellence
ORR	Overall response
OS	Overall survival
PBAC	Pharmaceutical Benefits Advisory Committee
PD	Progressive disease
PFS	Progression free survival
PR	Partial response

Abbreviation	Full definition
PRO	Patient reported outcomes
RCC	Renal cell carcinoma
RCO	Randomised cross over
RCTs	Randomised control trials
RDD	Randomised discontinuation design
RECIST criteria	Response Evaluation Criteria In Solid Tumours criteria
RMME	Repeated measures mixed-effect model
RR	Relative Risk
SD	Standard deviation
SE	Standard error
SMC	Scottish Medicines Consortium
SQL	Structured query language
SRDB	Heron Systematic Review Database
STA	Single technology appraisal
SU011248	Sunitinib
TARGET	Treatment Approaches in Renal Cancer Global Evaluation Trial
TIW	Three times a week
TTF	Time to failure
TTP	Time to progression
TTR	Time to response
UK	United Kingdom
UKCTG	UK Clinical Trials Gateway
US	United States
VEGF TKI	Vascular Endothelial Growth Factor- Tyrosine Kinase Inhibitor
WHO criteria	World Health Organisation criteria

Map of NICE manufacturers template to report sections

NICE manufacturer's template section	Report sections
5.1.1	3.1.2, 3.1.3
5.2.1	3.1.3
5.2.2	4.1
5.2.3	4.2
5.2.4	4.2
5.2.5	4.2
5.2.6	4.2
5.2.7	4.2
5.3.2	4.4.1
5.3.3	4.4.2
5.3.4	4.4.2
5.3.5	4.4.3
5.3.6	4.4.4
5.3.7	4.4.5
5.3.8	4.4.6
5.4	4.5
5.5	5
5.6	3.1.6, 6
5.7	3.1.2, 3.1.3, 3.1.7, 4.5, 7
5.8	4.2
5.9	3.1.2, 3.1.3, 5.3, 6.3, 7.3
6.1	3.2, 8
Appendix 2	Appendix B.1
Appendix 3	Appendix E
Appendix 4	Appendix B.1
Appendix 5	Appendix E
Appendix 8	Appendix B.1
Appendix 9	Appendix E
Appendix 10	Appendix B.2
Appendix 11	Appendix F

Executive summary

1.1 Objectives

Two reviews were conducted in order to address three objectives.

Primary objectives:

- To determine the relative clinical efficacy, safety and tolerability of pazopanib and other pharmacological interventions in the treatment of advanced/metastatic renal cell carcinoma in treatment-naïve patients.
- To determine the cost-effectiveness of pharmacological interventions in the treatment of advanced/metastatic renal cell carcinoma (RCC) in treatment-naïve patients.

Secondary objective:

- To determine the relative impact of pazopanib and other pharmacological interventions on quality of life in treatment naïve patients with advanced/metastatic renal cell carcinoma.

1.2 Data Sources

Comprehensive literature searches were conducted for the two reviews.

Clinical review:

- Examined the major literature databases MEDLINE, EMBASE and Cochrane (CENTRAL, CDSR and Methodology register).
- Timeframe of search was from database start to 23rd November 2009.
- Conference proceedings were searched in order to ensure all relevant literature was identified. Four conferences, ASCO, ASCO-GU, ESMO and ECCO were searched from 2007 to 2009.
- The reference lists of previous trials and systematic reviews, trials in progress databases (clinicaltrials.gov and UKCTG and ISRCTN) and MEDLINE in process were also searched.
- In order to provide a complete understanding of evidence regarding pazopanib versus current treatment for RCC, unpublished data from clinical study reports (CSRs) held by the manufacturer/sponsor were also included. The manufacturers of other technologies were not contacted for unpublished data.

Economic review

- The major databases for economic data were searched including MEDLINE, EMBASE and Cochrane (NHSEED, DARE, Technology Assessments database)
- Timeframe of search was from 1980 up to 23rd November 2009.
- In addition, supplementary searches of published health technology appraisals were conducted for the following authorities: NICE, SMC, CADTH, PBAC, AWMSG).

1.3 Review Methods

Clinical review:

- Comprehensive searches were run to identify studies which were potentially relevant to the review.
- To be included in the review trials had to meet pre-defined eligibility criteria; randomised controlled trials in treatment-naïve patients with advanced/metastatic RCC being treated with pazopanib, sunitinib, sorafenib, bevacizumab plus interferon (IFN), temsirolimus, IFN or interleukin-2 (IL-2) compared to placebo, BSC or IFN/IL-2.
- Studies were to have a full-text English publication to be included.
- Abstracts of citations found through the searches were initially reviewed for inclusion based on the abstract alone. Full-text copies of studies which potentially met the inclusion criteria or where it was not possible to determine whether the study could meet the inclusion criteria were ordered.
- Following receipt of all full-text papers, the eligibility criteria were applied to this full-text. At both stages, screening was conducted by two independent reviewers, and any discrepancies between reviewers were reconciled by a third independent reviewer.
- Studies which met the eligibility criteria after the second screening stage were extracted in parallel by two independent reviewers and any discrepancies reconciled by a third party.
- Where more than one publication was identified describing a single trial, the data were compiled into a single entry in the data extraction table. Both quantitative and qualitative (meta-analysis and indirect) analyses were conducted on data extracted from the included studies.

Economic review:

- Comprehensive searches were run to identify studies which were potentially relevant to the review.
- To be included in the review trials had to meet pre-defined eligibility criteria; an economic evaluation of interventions (the same as those in the clinical review) in treatment naïve patients with advanced/metastatic RCC.
- Studies were to have a full-text English publication to be included.
- The abstracts of citations retrieved by the searches were screened in the same fashion as for the clinical review. Additionally, studies which met the eligibility criteria after the second screening stage were also extracted in the same manner to the clinical review.

1.4 Results

Clinical Review

Studies identified

Thirteen trials meeting the clinical review inclusion criteria were identified in the review process and contributed to the qualitative and quantitative analysis.

No studies which directly compared pazopanib, the intervention under consideration, with one of the other interventions of interest (sunitinib, sorafenib, bevacizumab plus IFN,

temsirolimus, IFN, IL-2) were identified. One trial that examined the clinical efficacy and safety of pazopanib compared to placebo was identified. A second pazopanib study (VEG102616) was found which had a randomisation discontinuation design; however, owing to the fact that this trial was revised to be treated like a single-arm open-label trial and a lack of outcome data for the treatment-naïve subgroup for the randomised phase of the trial, this was excluded from the final list of included studies. In addition, 12 studies comparing the other interventions of interest to placebo or to IFN were identified, permitting indirect comparisons of the interventions.

Methodology

From the results of direct meta-analyses, network diagrams were produced for each efficacy and safety outcomes extracted, to identify the possible indirect analyses which could be performed. The inputs in the indirect analyses were the results from direct analyses using the random effects model. This is due to the additional heterogeneity that exists in such analyses, which renders the assumptions underlying a fixed effects model less reasonable. Indirect analyses were conducted using all data available for each outcome in the IFN versus BSC trials. However, additional indirect analyses were conducted using only the MRC RE01 trial for the IFN versus BSC comparison.

Efficacy outcomes

The qualitative and quantitative analyses showed that pazopanib had superior efficacy compared to placebo and comparable efficacy to sunitinib, sorafenib and bevacizumab plus IFN. No significant differences between pazopanib and the comparators for PFS and OS were seen in the indirect analyses. Similarly, there were no significant differences between pazopanib and the comparators in response rates, although all treatments demonstrated significantly higher overall response rates than placebo/BSC and/or IFN.

The relative efficacy of pazopanib and temsirolimus could not be determined owing to significant differences in the trial populations. Although there was an overlap in the patient populations in the two trials, a much higher proportion of patients in the temsirolimus trial had a poorer prognosis than those enrolled in the pazopanib trial in terms of their Memorial Sloan Kettering Cancer Centre (MSKCC) risk score.

Safety outcomes

As expected owing to the nature of the treatment modalities, all active interventions were associated with a high proportion of patients (generally >90%) experiencing an adverse event (AE). Pazopanib showed a lower risk of patients experiencing any AE and any grade 3 or 4 AE compared to sorafenib and bevacizumab plus IFN; however the difference was not statistically significant. A comparison with sunitinib could not be carried out in this respect since the percentage of patients experiencing any AE or any grade 3 or 4 AE was not reported in the sunitinib trial publication.

When examining specific individual AEs, the safety profile for pazopanib was favourable compared to sunitinib, sorafenib, temsirolimus and bevacizumab plus IFN, particularly for AEs in the class of blood and lymphatic disorders including anaemia. Pazopanib was associated with a reduced risk of almost all AEs for which comparisons could be conducted compared to sunitinib. These included diarrhoea, vomiting, fatigue, hand-foot syndrome, total bilirubin increased, anorexia, epistaxis, and haematological AEs. The differences in risk, however, rarely reached statistical significance.

Indirect analysis for safety profile across all treatments was performed. Through this indirect analysis, pazopanib demonstrated comparable, and for certain AEs, improved safety. This was particularly evident for the blood and lymphatic disorder AEs such as anaemia, where the risk of these events was consistently lower with pazopanib than the

comparators, albeit these differences did not reach statistical significance. The qualitative findings also suggested that pazopanib had an improved safety profile, particularly with regard to haematological side effects, over sunitinib. The exceptions in improved safety profile were hair colour change and AST and ALT levels. A greater proportion of patients with increased levels of AST and ALT were observed with pazopanib as compared to sunitinib.

Again, specific AEs experienced by patients treated with the active interventions were generally mild to moderate; relatively few patients (<25%) suffering grade 3 or 4 AEs. Additionally, for many of the specific AEs, such as headache, no patients on pazopanib experienced a grade 3 or 4 AE, suggesting improved safety profile over the comparators.

Data limitations

The evidence base would benefit from head-to-head comparisons, particularly those including sunitinib, since it would provide more robust conclusions for some outcomes where indirect analysis has produced inconclusive results. It would also benefit from further comparisons of safety outcomes between treatments, since the older IFN trials, which are key to the indirect comparison of pazopanib to other treatments, do not report some AE outcomes which have been a more recent focus.

Economic review

Studies identified

The economic review identified only two studies which met the inclusion criteria for the review. One was a Markov Model based study assessing the cost effectiveness and cost utility of sunitinib as a first-line treatment in metastatic RCC compared with IFN and IL-2 from a US societal perspective (Remak 2008). The second was a decision analytical model based study evaluating the costs of managing AEs of bevacizumab plus IFN compared to sunitinib in the first-line treatment of metastatic RCC in United Kingdom, Germany, Italy and France (Mickisch 2009).

Results

The results of the economic evaluation presented by Remak et al (Remak 2008) showed that sunitinib was both less costly and more effective than IL-2. In addition, sunitinib was more costly, but more effective than IFN, resulting in an ICER (life years gained) of $67 215 and an ICUR of $52 593. Mickisch 2009 reported that the average cost per patient of managing all-grade and grade-3-4 AEs varied across the countries assessed in the evaluation, and that the costs were higher for sunitinib than for bevacizumab plus IFN. The main cost drivers were lymphopenia, neutropenia, thrombocytopenia, leucopenia and fatigue/asthenia for sunitinib; and proteinuria, fatigue/asthenia, bleeding, anaemia and gastrointestinal perforation for bevacizumab plus IFN.

1.5 Conclusion

Pazopanib is an effective treatment for treatment-naïve patients with advanced/metastatic renal cell carcinoma, demonstrating superior efficacy compared to BSC and comparable efficacy to the current standard of care, sunitinib. Additionally, pazopanib was also associated with an improved safety profile, particularly for haematological AEs, over sunitinib.

2 Introduction

2.1 Aims of the Review

The purpose of this review is to provide evidence that will support the efficacy and safety of pazopanib in the treatment of advanced/metastatic renal cell carcinoma (RCC) in treatment-naïve patients. The systematic review was undertaken and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines.

The primary study question addressed by this review is:

- What is the relative clinical efficacy, safety and tolerability of pazopanib and other pharmacological interventions in the treatment of advanced/metastatic RCC in treatment-naïve patients?
- What is the cost-effectiveness of pharmacological interventions in the treatment of advanced/metastatic RCC in treatment-naïve patients?

The secondary study question addressed by this review is:

- What is the effect of pazopanib and other pharmacological interventions on quality of life endpoints in advanced/metastatic RCC in treatment-naïve patients?

2.2 Background

2.2.1 Aetiology and epidemiology

Renal cell carcinoma (RCC) is an aggressive kidney cancer which originates in the renal parenchyma and accounts for approximately 90% of kidney cancers and is responsible for almost 3% of all adult cancers (NICE 2009b). It is almost twice as common in men as women, and as with many other cancers, the risk of developing the disease increases with age. The incidence of RCC is highest in people older than 65 years of age (NICE 2009b). In 2006, there were 7840 new cases of kidney cancer in the United Kingdom (UK), and 3752 deaths in 2007 from the disease (cancerresearch.co.uk).

Several risk factors have been identified for RCC and include:

- Smoking
- Obesity
- Hypertension
- Von Hippel-Lindau disease (Pascual 2008).

2.2.2 Clinical presentation of RCC

In its early stages, RCC is usually asymptomatic or has only mild symptoms, and affected individuals are often diagnosed incidentally as a result of imaging performed for unrelated reasons (Larkin 2009). Symptoms of RCC appear when the cancer enlarges or spreads to other parts of the body; approximately a third of patients present with advanced and/or metastatic disease (Larkin 2009; Harrison 2007).

Symptoms of RCC may include the following:

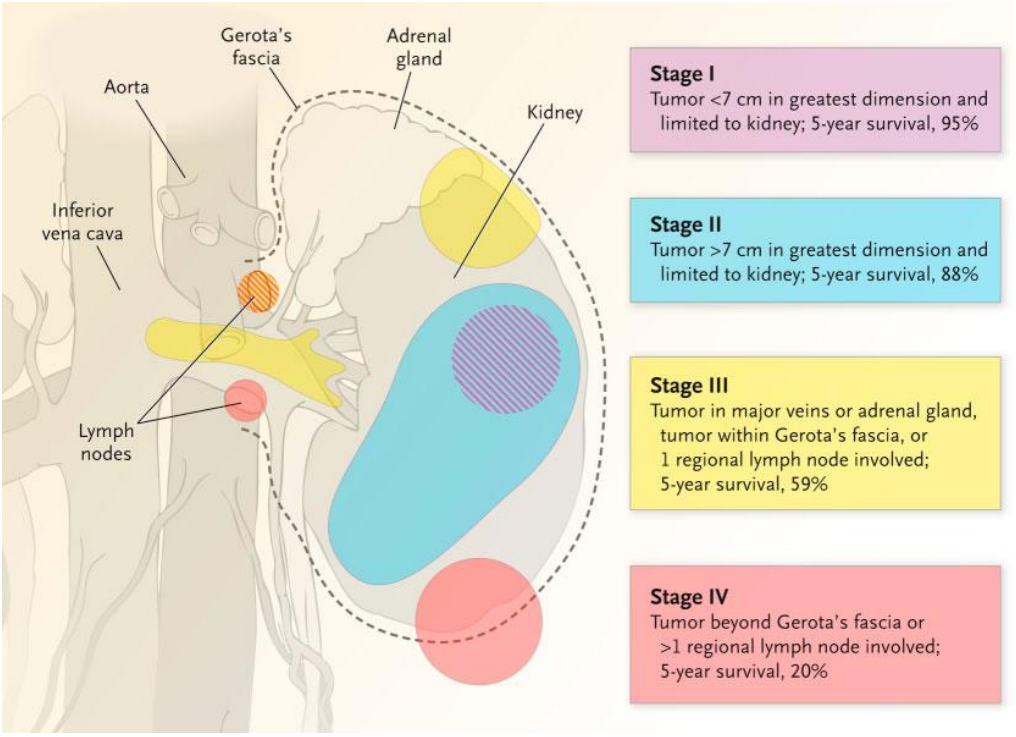
- Haematuria (blood in the urine)
- Abdominal pain
- Palpable abdominal mass
- Loss of appetite
- Weight loss
- Pyrexia (fever)
- Night sweats
- Malaise
- Anaemia

RCC can also cause a number of paraneoplastic syndromes, which are conditions caused by the tumour when it releases cytokines or hormones. Symptoms of these conditions include hypertension and hypercalcaemia.

2.2.3 Staging of RCC

One of the determining factors for the treatment of RCC, as with many cancers, is the stage of the disease at diagnosis. Therefore following initial diagnosis of the disease, cases are confirmed through a series of procedures and laboratory tests. The patient's prognosis is also closely related to the stage of the disease. Figure 1 shows the criteria for each stage of the disease.

Figure 1: Staging of RCC (Cohen 2005)



The Memorial Sloan-Kettering Cancer Centre (MSKCC) has produced a predictive tool for survival based on five variables, namely low Karnofsky performance status (< 80%), high lactate dehydrogenase (>1.5 times upper limit of normal), low serum haemoglobin (< lower limit of normal), high corrected serum calcium (>10 mg/dL), and absence of prior nephrectomy. Patients are then assigned to one of three risk groups according to the number of risk factors they exhibit: three-year survival percentages for the favourable-

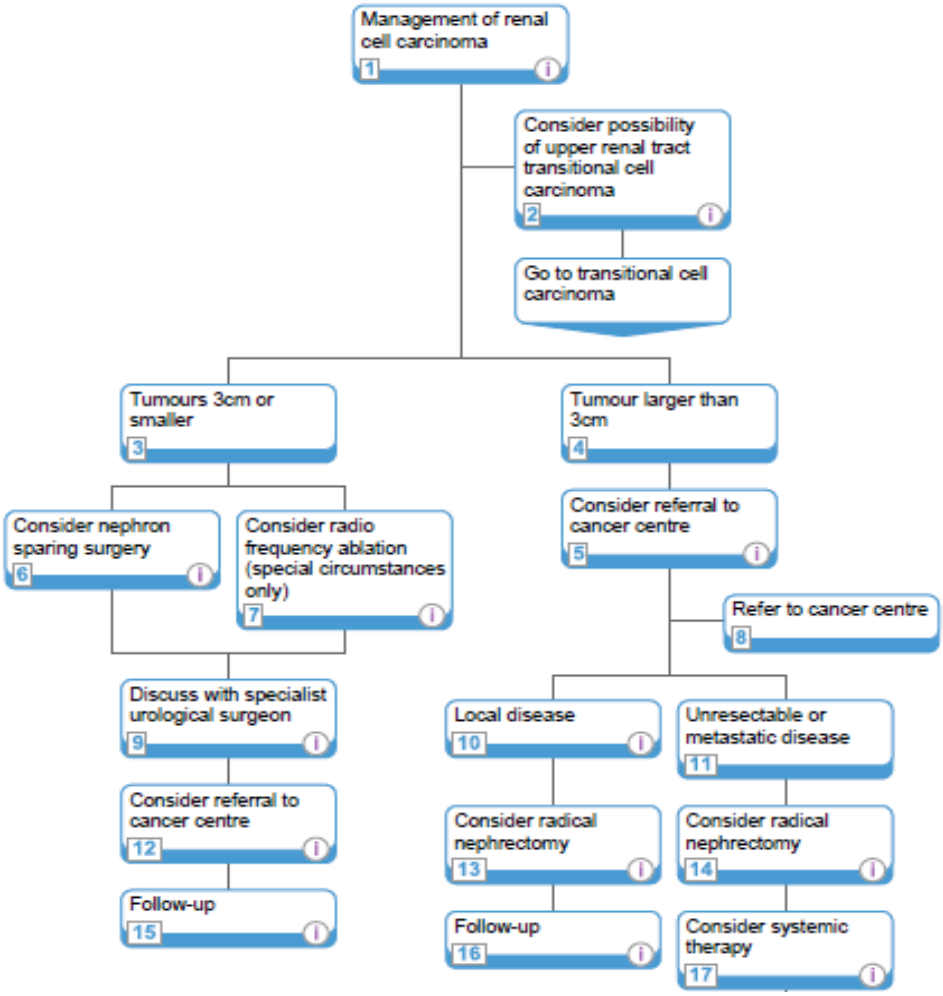
risk (no risk factors), intermediate risk (one or two risk factors), and poor-risk (three or more risk factors) groups were 31%, 7%, and 0%, respectively (Motzer 1999).

2.2.4 Current treatment guidance and practice

Treatment of RCC is dependent on a number of factors including the stage of disease and the health of the patient. Other factors which are considered while selecting the treatment include: type of RCC, the size of tumour (and its location) and the age of the patient.

Surgery is often the primary treatment modality for early stage/localised disease, whereby either the whole or part of the kidney, are removed in a procedure called nephrectomy (Figure 2). However, up to 30% of patients who undergo curative surgery for localised disease relapse and develop metastatic disease (BMJ Evidence Centre 2010). There are currently no treatments that reliably cure advanced and/or metastatic RCC (NICE 2009b). It is one of the most difficult-to-treat malignancies being largely resistant to chemotherapy, radiation therapy and hormonal therapy.

Figure 2: NHS treatment pathway (http://healthguides.mapofmedicine.com/choices/map/index.html)



Until recently, cytokines (IFN, IL-2), categorised under immunotherapy, were the only available treatments. However, their use has been limited by their modest response rates and significant toxicity (Harrison 2007; Athar 2008; Garcia 2007). Immunotherapy with IFN and/or IL-2 has demonstrated response rates of approximately 10-20%, with few

studies showing response rates above 20%, and some demonstrating rates as low as 6% (Harrison 2007). Further, some benefit may be attributed to prior nephrectomy, since response rates were observed to be higher in patients who underwent nephrectomy prior to IFN (Garcia 2007). Additionally the majority of responses are short lived and often do not last more than one year (Athar 2008). Significant toxicity, particularly with the use of IL-2 is also observed and can result in treatment delays and reduction in dose (Athar 2008).

Understanding of the biology of renal cell cancer has increased in recent years leading to a rationale target based approach to treatment. Von-Hippel Lindau (VHL) disease is characterised by an increased risk of developing renal cell cancer. The VHL gene encodes a 213 amino acid protein (pVHL). When VHL gene functions normally, pVHL is the substrate recognition component of an ubiquitin ligase complex. This complex targets hypoxia-inducible factor (HIF) (a protein transcription factor) for proteolysis. When the conditions are hypoxic or when mutation of the VHL gene occurs, the interaction between pVHL and HIF gets disturbed and thus HIF is constitutively activated. Translocation of HIF into the nucleus takes place leading to transcription of hypoxia-inducible genes, including VEGF and PDGF. Examinations of RCC tumour samples have shown VEGF overexpression that drives angiogenesis. Thus, VEGF inhibition has become a therapeutic target in RCC (Cockman 2000; Iliopoulos 1996).

The introduction of therapies which target VEGF receptors has greatly impacted the management of this disease area, with significant clinical activity demonstrated in both treatment-naïve and cytokine pre-treated patients. However, only sunitinib has been recommended by the National Institute for Health and Clinical Excellence (NICE) for the first-line treatment of advanced/metastatic RCC (NICE 2009a). No therapies are recommended by NICE as second-line treatment options (NICE 2009b). Table 1 lists those therapies currently licensed or with a license pending for the treatment of advanced/metastatic RCC in the UK.

Table 1: Summary of targeted agents for the treatment of advanced/metastatic RCC (http://www.emc.medicines.org.uk/ accessed on 2 February 2010)

Treatment	Brand name	Mechanism of action	Indication	Manufacturer	Date of licence (EMEA)	NICE decision
First-line						
Pazopanib	Votrient®	Selectively inhibits vascular endothelial growth factor receptors (VEGFR)-1, -2 and -3, c-kit and platelet derived growth factor receptors (PDGF-R) α and β, which may result in inhibition of angiogenesis in tumours in which these receptors are upregulated. Pazopanib has minimal activity on Fms-related tyrosine kinase 3 (FLT3).	**Anticipated licence – First-line** treatment for advanced renal cell carcinoma (RCC) and in patients that have received cytokine treatment for advanced disease	GSK	-	STA in progress
Sunitinib	Sutent®	Blocks the tyrosine kinase activities of vascular endothelial growth factor receptor (VEGFR) 1, 2, 3 , platelet-derived growth factor receptor(PDGFR) α and β, colony stimulating factor receptor type 1 (CSF-1R), glial cell-line derived neurotrophic factor receptor (RET) and c-kit, thereby inhibiting angiogenesis and cell proliferation. This agent also inhibits the phosphorylation of Fms-related tyrosine kinase 3 (FLT3), another receptor tyrosine kinase expressed by haematological progenitor cells.	Treatment of advanced and/or metastatic RCC	Pfizer	July 2006	Recommended (March 2009)
Bevacizumab	Avastin®	Binds to VEGF and inhibits VEGF binding to the VEGF receptor, thereby preventing the growth and maintenance of tumour blood vessels.	In combination with IFN; first-line treatment of patients with advanced and/or metastatic renal cell cancer	Roche	January 2005	Not recommended (August 2009)
Temsirolimus	Torisel®	Binds to and inhibits the mammalian target of rapamycin (mTOR), resulting in decreased expression of mRNAs necessary for cell cycle progression and arresting cells in the G1 phase of the cell cycle. mTOR is a serine/threonine kinase which plays a role in the PI3K/AKT pathway that is upregulated in some tumours.	First-line treatment of patients with advanced renal cell carcinoma (RCC) who have at least three of six prognostic risk factors	Pfizer	November 2007	Not recommended (August 2009)
Second-line						
Sorafenib	Nexavar®	Blocks the enzyme RAF kinase, a critical component of the RAF/MEK/ERK signaling pathway that controls cell division and proliferation; in addition, sorafenib inhibits the VEGFR-2/PDGFR-beta signaling cascade, thereby blocking tumour angiogenesis.	Treatment of patients with advanced renal cell carcinoma who have failed prior IFN or interleukin-2 based therapy or are considered unsuitable for such therapy	Bayer	July 2006	Not recommended (August 2009)
Everolimus	Afinitor®	Everolimus is a selective mTOR (mammalian target of rapamycin) inhibitor, resulting in interference with the translation and synthesis of proteins. It reduces levels of vascular endothelial growth factor (VEGF), which potentiates tumour angiogenic processes.	Treatment of patients with advanced renal cell carcinoma, whose disease has progressed on or after treatment with VEGF-targeted therapy.	Novartis	August 2009	Not recommended (February 2010)

EMEA = European Medicines Agency, GSK= GlaxoSmithKline, mTOR= Mammalian target of rapamycin, NICE= National Institute for Health and Clinical Excellence.

2.3 Rationale for the Review

Pazopanib, an oral multi-kinase inhibitor, is currently undergoing regulatory review in Europe for the treatment of advanced/metastatic RCC. Pazopanib is also being considered within NICE's work programme and a Single Technology Appraisal (STA) for the use of pazopanib in the first-line treatment of advanced/metastatic RCC is in progress. The primary aim of this review is therefore to determine the relative clinical efficacy, safety, tolerability and quality of life of pazopanib and other pharmacological interventions in the first-line treatment of advanced/metastatic RCC.

3 Review methods

3.1 Clinical systematic review

The systematic review was undertaken and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines (Moher 2009).

3.1.1 Study protocol

A full protocol for the systematic review was written detailing the patient population, interventions and study designs to be analysed. A summary version of the original study protocol can be found in Appendix A.1.

3.1.2 Identification of studies

A comprehensive search strategy was designed to retrieve relevant clinical data from published literature; details of the search strategy can be seen in Appendix B. In addition to the database searches conference searching was also conducted to ensure all relevant literature was included in the review.

Trials in progress were also identified, in order to highlight future studies which may be published and to provide additional data for the treatments of interest. Details of the search methodologies are provided in Appendix B.

The following databases were examined from 1980 up to 23 November 2009, Table 2.

Table 2: Databases examined for the clinical systematic review and the service provider used

Data source	Service Provider
MEDLINE	Embase.com; http://www.embase.com/
EMBASE	
Cochrane Central Register of Controlled Trials (CENTRAL)	Cochrane library; http://mrw.interscience.wiley.com/cochrane/cochrane_search_fs.html
Cochrane Database of Systematic Reviews (CDSR)	
Cochrane Methodology Register	
MEDLINE in process (2009 only)	PubMed; http://www.ncbi.nlm.nih.gov/sites/entrez

The following conference proceedings were hand searched from 2007 to 2009:

- American Society of Clinical Oncology (ASCO)
- ASCO-Genitourinary (ASCO-GU)
- European Society for Medical Oncology (ESMO)
- European Conference for Clinical Oncology (ECCO)

Other data sources:

- Bibliographic searching of previous trials and systematic reviews
- Clinicaltrials.gov (02 December 2009)
- UK clinical trials gateway (UKCTG) and International Standard Randomised Controlled Trial Number (ISRCTN) (02 December 2009)

In order to provide a complete understanding of the evidence for pazopanib compared with current treatments for advanced/metastatic RCC, additional sources of information other than those identified by the systematic review have been used to supplement data

in sections 4 to 7; unpublished data from clinical study reports (CSRs) held by the manufacturer/sponsor were also included. The manufacturers of other technologies were not contacted for unpublished data.

3.1.3 Study selection

The list of eligibility criteria is provided in Table 3. The outcomes which were used to determine inclusion in the qualitative and quantitative analysis are listed in section 3.1.6.

Table 3: Eligibility criteria

	Clinical effectiveness	Rationale
Inclusion criteria	**Population** • Age: Adults (≥ 18 years) • Gender: Any • Race: Any • Stage of disease: Advanced and Metastatic (stage III/IV) • Line of therapy: Treatment naïve	• The patient population has been restricted to match that stated in the decision problem for pazopanib in the first-line treatment of advanced/metastatic RCC. • Since the current treatments for RCC are licensed for adult patients, studies including children or adolescents were excluded.
	Interventions • Pazopanib monotherapy (or in combination with BSC) • IFN-α monotherapy (or in combination with BSC) • IL-2 monotherapy (or in combination with BSC) • Sunitinib monotherapy (or in combination with BSC) • Sorafenib monotherapy (or in combination with BSC) • Temsirolimus monotherapy (or in combination with BSC) • Bevacizumab in combination with IFN-α (and in combination with BSC)	• The included interventions are those which are either licensed for the first-line treatment of advanced/metastatic RCC or for which RCT data in this setting exist. • The review was limited to studies of these agents administered as monotherapy (or with the exception of bevacizumab in combination with IFN) as per their licensed indications or as per the anticipated licence in the case of pazopanib.
	Comparator • A different intervention from the included list • Placebo • Best supportive care (BSC)*	• These comparators were chosen to enable both direct and indirect comparisons between the interventions of interest.
	Study design • Randomised control trials (RCTs) with any blinding status	• RCTs are the gold standard of clinical evidence, minimising the risk of confounding and allowing the comparison of the relative efficacy of interventions. Therefore only these studies were included. To enhance the level of evidence, studies with double blind, single blind and open label design were included.
	Language restrictions • English only	• The restriction would not limit results substantially due to data availability in English language.
	Publication timeframe • 1980 onwards for literature searches • Last 3 years for conference searching	• Studies which are presented at conference are usually published in journals within 3 years.

	Outcome of interest
	• Studies should report an outcome of interest.
	• Outcomes of interest are:
	  ◦ Overall and progression-free survival
	  ◦ Complete, partial, overall response and stable disease
	  ◦ Time to response
	  ◦ Duration of and time to response
	  ◦ Withdrawals
	  ◦ Quality of life
	  ◦ Safety
		• Studies which do not report outcomes of interest would not feature in any analyses and were therefore were excluded.
		• These outcomes were chosen since they are frequently measured and reported in trials of RCC, and will enable the study question of the review to be answered. Moreover, these outcomes were also referred from the NICE submission of sunitinib.

Exclusion criteria	**No subgroup analysis** • No subgroup analysis for disease of interest • No subgroup analysis for advanced/metastatic disease • No subgroup analysis for treatment naïve patients	• Studies with no subgroup data for the disease, disease stage and line of treatment were not included, since these studies would introduce heterogeneity into the review.

*BSC definition: no active treatment/observation/a method of care that is not a focused treatment/treatments which clinicians consider to be "placebo-equivalent" including medroxyprogresterone and vinblastine. RCC= renal cell carcinoma , RCT = Randomised control trial

Further detail regarding the inclusion criteria for the intervention and comparator and the rationale for this are provided in Table 4.

Table 4: Detail regarding the intervention and comparator inclusion criteria

Intervention	Comparator	Included?	Reasoning
Intervention list drug (e.g. Sunitinib)	BSC/Placebo	Yes	Allow us to obtain studies with common comparators which can then be indirectly compared.

Intervention	Comparator	Included?	Reasoning
Intervention list drug (e.g. Sunitinib)	Other intervention list drug (e.g. Pazopanib)	Yes	This allows direct comparison of interventions.
Intervention list drug (e.g. Sunitinib)	Non-intervention list drug (e.g. Surgery)	No	These studies do not aid in answering the study question, and would not provide useful links in indirect comparisons.
Intervention list drug + Non-intervention list drug (e.g. Sunitinib + Retinoic acid)	Non-intervention list drug (e.g. Surgery)	No	
Intervention list drug dose 1 (e.g. IFN dose 1)	Intervention list drug dose 2 (e.g. IFN dose 2)	No	
Intervention list drug A + intervention list drug B (e.g. Sunitinib + IFN)	Intervention list drug C (e.g. Pazopanib)	No	
Intervention list drug A + intervention list drug B (e.g. Sunitinib + IFN)	Intervention list drug A (e.g. Sunitinib)	No	
Intervention list drug A + non-intervention list drug (e.g. IFN + retinoic acid)	Intervention list drug B (e.g. IL-2)	No	
Intervention list drug A + non-intervention list drug (e.g. IFN + Retinoic acid)	Intervention list drug A (e.g. IFN)	No	

3.1.4 Data Extraction Strategy

Bibliographic details and abstracts of all citations detected by the literature search were downloaded into the Heron Systematic Review Database (SRDB), a bespoke, structured query language (SQL)-based internet database.

First pass of citations

Citations were first screened based on the abstract supplied with each citation. Each citation was screened by two independent reviewers, and any discrepancies between reviewers were reconciled by a third independent reviewer. Citations that did not match the eligibility criteria were excluded at this 'first pass'; where unclear, citations were included. Duplicates of citations (due to overlap in the coverage of the databases) were also excluded in the first pass. Full-text copies of all references that could potentially meet the eligibility criteria were also ordered at this stage.

Second pass of citations

The eligibility criteria were applied to the full-text citations. Each full-text was screened by two independent reviewers, and any discrepancies between reviewers were reconciled by a third independent reviewer. Data presented in the studies still included after this stage were extracted to data extraction grids.

Extraction strategy

The final extraction grid is provided in Appendix C.1. Data from trials were extracted in parallel by two independent reviewers, with reconciliation of any differences by a third independent reviewer. All reviewers were qualified with either a Masters degree in pharmacy or an equivalent related discipline, and furthermore were fully trained in conducting systematic reviews with a minimum of 1.5 to 2 years of full-time experience of systematic review work within a health economics and outcomes research organisation.

Where more than one publication was identified describing a single trial, the data were compiled into a single entry in the data extraction table to avoid double counting of patients. Each publication was referenced in the table to recognize that more than one publication may have contributed to the entry.

Studies excluded during data each stage, along with rationale for exclusion are provided in a separate MS Excel document (Clinical Excluded Studies).

3.1.5 Quality Assessment

Studies were assessed qualitatively and by means of a study grade and Jadad score. In addition, a qualitative assessment was conducted, using the assessment criteria recommended in the NICE manufacturer's template.

3.1.5.1 Qualitative Assessment

A descriptive analysis of each extracted study was made during the data extraction process. The analysis assessed the study for quality by considering the following features, which could introduce bias (Table 5).

Table 5: Criteria for qualitative assessment

Criterion	Assessment
Randomisation	Was randomisation carried out appropriately?
Allocation concealment	According to the grading system (see section 3.1.5.2)
Baseline comparability:	Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?
Blinding	Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?
Follow-up	Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?
Selective reporting	Is there any evidence to suggest that the authors measured more outcomes than they reported?
Analysis	Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

3.1.5.2 Study Grading

In addition to qualitative assessment, study quality was also graded according to two scales. The first assessed the adequacy of concealment of allocation. The second was the Jadad score, which assesses study quality and study reporting (Jadad 1996).

Concealment of allocation

Concealment of allocation was graded for each study, where:

(A) Allocation concealment was adequate (e.g. centralized allocation by a central office unaware of subject characteristics; pharmacy-controlled randomisation; pre-numbered or coded identical containers which are administered serially to participants; on-site computer system combined with allocations kept in a locked unreadable computer file that can be accessed only after the characteristics of an enrolled participant have been entered; sequentially numbered, sealed, opaque envelopes.)

(B) Unclear (when studies do not report any concealment approach, adequacy should be considered unclear).

(C) Inadequate (e.g. the use of case record numbers, dates of birth or day of the week, and any procedure that is entirely transparent before allocation, such as an open list of random numbers.)

(D) Allocation concealment was not used.

Jadad score

One point was scored for each positive answer to the following questions:

1. Was the study randomised?
2. Were the randomisation methods used adequate?
3. Was the study described as double-blind?
4. Were blinding methods adequate?
5. Was there a description of withdrawals and drop-outs?

Points were subtracted if randomisation or blinding methods were judged to be obviously flawed.

3.1.6 Statistical analysis

The general statistical methods used in the meta-analyses are described in Appendix F. Comparison of efficacy, safety and tolerability outcomes were made between interventions by pooling data from studies using standard meta-analytic techniques. A relative effect between two treatments was also determined using the same method, where only one study provided data for a particular outcome.

Interventions were grouped according to Table 6.

Table 6: Grouping of interventions for the meta-analysis

Intervention group	List of interventions within group
Pazopanib	Pazopanib monotherapy, Pazopanib + BSC/Placebo
Sunitinib	Sunitinib monotherapy, Sunitinib + BSC/Placebo
Sorafenib	Sorafenib monotherapy, Sorafenib + BSC/Placebo
Temsirolimus	Temsirolimus monotherapy, Temsirolimus + BSC/Placebo
Bev + IFN	Bevacizumab plus IFN, Bevacizumab plus IFN + BSC/Placebo
IFN	IFN monotherapy, IFN + BSC/Placebo
IL-2	IL-2 monotherapy, IL-2 + BSC/Placebo
Placebo	Placebo, BSC
BSC	No active treatment, Medroxyprogesterone, Vinblastine

Bev = Bevacizumab, BSC = best supportive care, IL2 = interleukin-2

Outcomes selected for the review were referred from the NICE submission of sunitinib. The outcomes included in the analysis were those in Table 7. Further rationale for the inclusion of these outcomes is presented in Table 3.

Table 7: Outcomes included in the review

Outcome group	Type of data	Specific outcome
Efficacy	Dichotomous	Overall survival (OS) at 1 year
		OS at 2 years
		OS at endpoint
		Progression free survival (PFS) at 1 year
		PFS at 2 years
		PFS at endpoint
		Complete response (CR)
		Partial response (PR)
		Overall response
		Stable disease
	Continuous	OS (Intention to treat (ITT) and cross over adjusted)
		PFS
		Time to progression (TTP)
		Duration of response (DOR)
		Time to response (TTR)
Tolerability	Dichotomous	Withdrawal due to AEs

Outcome group	Type of data	Specific outcome
		Withdrawal due to death
		Total withdrawals
		Down dosing
		Dose interruption
Quality of life	Continuous	FACT-Kidney Symptom Index–Disease-related Symptom subscale (FKSI-DRS Index)
		FACT-Kidney Symptom Index–15 item scale (FKSI-15 Index)
		Functional assessment of cancer therapy – general scale (FACT-G)
		EuroQoL Questionnaire (EQ-5D)
		EuroQoL -Visual Analog Scale (EQ-VAS)
		EORTC Quality of Life-Questionnaire-C30 (EORTC QLQ C-30)
Safety	Dichotomous	Any AE (all grades)
		Any grade 3 or 4 AE
		Any serious AEs
		Any treatment related AEs
		Abdominal pain
		Alopecia
		Alanine transaminase (ALT) increased
		Altered taste
		Anaemia
		Anorexia
		Arthralgia
		Aspartate transaminase (AST) increased
		Asthenia
		Congestive heart failure (CHF)
		Depression
		Diarrhoea
		Dyspepsia
		Epistaxis
		Fatigue
		Fever
		Flank pain
		Flu-like symptoms
		Hair colour change
		Hand-foot syndrome
		Headache
		Hyperglycaemia
		Hypertension
		Hypoglycaemia
		Hypothyroidism
		Hypophosphataemia
		Infection
		Leucopenia
		Lymphocytopaenia
		Mucositis/stomatitis
		Nausea
		Neutropenia
		Rash
		Skin discolouration
		Thrombocytopenia
		Total bilirubin increased
		Vomiting

AE = Adverse event, ALT = Alanine transaminase, AST = Aspartate transaminase, CHF = congestive heart failure, DOR = Duration of Response, EORTC QLQ C-30 = EORTC Quality of Life-Questionnaire-C30, EQ-VAS = EuroQoL -Visual Analog Scale, EQ 5D = EuroQoL Questionnaire, FACT-G = Functional assessment of cancer therapy – general scale, FKSI-15 Index = FACT-Kidney Symptom Index–15 item scale, ITT = intention to treat, TTP = time to progression, TTR = time to response.

The minimum requirements for a meta-analysis to be conducted for a particular combination of outcome, treatment comparison and study grouping were as follows:

- For dichotomous and ordinal outcomes, the number of patients in which the outcome was observed and the intention to treat analysis (ITT) treatment arm population;

- For continuous outcomes, the mean and standard deviation (SD) for that outcome and the number of patients in which those statistics were calculated.

Stata statistical software was used, in particular the *metan* meta-analysis command (Harris 2007). Dichotomous outcomes were summarised as (relative) risk ratios and continuous outcomes were summarised as (weighted) mean differences (Egger 2001).

Summary statistics are presented with 95% confidence intervals (CIs) throughout this review. P values were also computed – these estimate the probability that the differences between the interventions could have arisen by chance due to sampling variability. Statistical assessment of heterogeneity was presented by means of I^2 statistics.

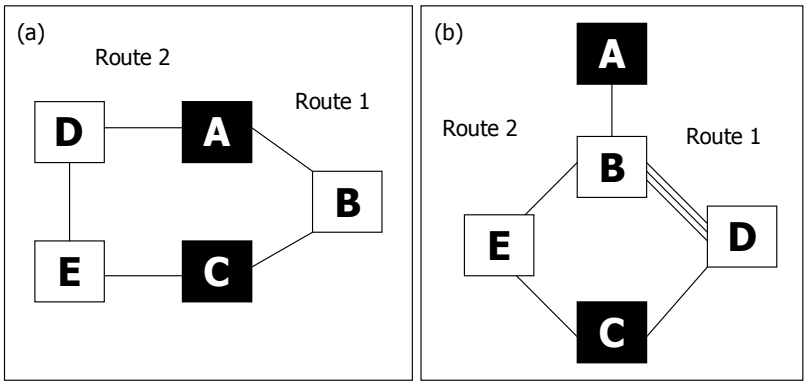
The results of the meta-analyses are displayed using 'forest plot' diagrams. These diagrams display the results from each study on a separate row, the pooled results appearing on the bottom row. The graphical display plots the treatment effect from each trial as a box with its CIs marked as a line extending either side of the box. The relative sizes of the boxes reflect the contribution each study makes to the overall summary. The overall summary statistic is given as a diamond shape, where the widest part in the centre is the point estimate and the horizontal width is the CI. Data are reported as mean ± SD unless otherwise specified.

3.1.7 Indirect analysis methodology

From the results of direct meta-analyses, network diagrams were produced for each efficacy and safety outcome extracted, to identify the possible indirect analyses which could be performed. The first choice of indirect comparison to perform was one via a single common comparator, for example, in Figure 3a, to compare treatment A with treatment C, route 1 was used. If there were no possible indirect comparisons via a single common comparator, comparisons were made via two comparators. If there was an option between two routes, a route with more trial results feeding into the meta-analysis was chosen. For example, in Figure 3b, to compare treatment A with treatment C route 1 would be used, where a connecting line equals one trial input into the meta-analysis.

When one route was preferable to another route, using the rules described above, only the first analysis was performed. When two routes were equal in preference, both analyses would be performed.

Figure 3: Network diagrams



(a) Compare treatment A with treatment C via common comparator B (route 1)

(b) Compare treatment A with treatment C via common comparators B and D (route 1) where each connecting line equals one trial input into the meta-analysis

The inputs in indirect analyses were the results from direct analyses using the random effects model. This is due to the additional heterogeneity that exists in such analyses, which renders the assumptions underlying a fixed effects model less reasonable. Indirect analyses were performed according to method described by Bucher et al (Bucher 1997).

3.2 Economic Review

3.2.1 Study protocol

A full protocol for the systematic review was written detailing the patient population, interventions and study designs to be included. A summary version of the original study protocol can be found in Appendix A.2.

3.2.2 Identification of studies

A comprehensive search strategy was designed to retrieve relevant economic data from published literature; details of the search strategy can be seen in Appendix B.

The following databases were examined from 1980 up to 23rd November 2009, Table 8:

Table 8: Databases examined for the economic systematic review and the service provider used

Data source	Service Provider
MEDLINE	Embase.com; http://www.embase.com/
EMBASE	
Cochrane Economic Evaluations Database/ NHS Economic Evaluation Database	Cochrane library; http://mrw.interscience.wiley.com/cochrane/cochrane_search_fs.html
Cochrane Technology Assessments Database	
Database of Abstracts of Reviews of Effects (DARE)	

DARE = Database of Abstracts of Reviews of Effects, NHS = National Health Service

In addition to the search of literature databases, supplementary searches of published health technology appraisals were conducted for the following authorities:

- NICE

- SMC
- CADTH
- PBAC
- AWMSG

3.2.3 Study selection

To be included in this study, trials were required to meet the eligibility criteria in Table 9.

Table 9: Eligibility criteria

	Clinical effectiveness	Rationale
Inclusion criteria	**Population** • Age: Adults (≥ 18 years) • Gender: Any • Race: Any • Stage of disease: Advanced and Metastatic (stage III/IV) • Line of therapy: Treatment naïve	• The patient population has been restricted to match that stated in the decision problem for pazopanib in the first-line treatment of advanced/metastatic RCC. • Since the current treatments for RCC are licensed for adult patients, studies including children or adolescents were excluded.
	Interventions • Pazopanib monotherapy (or in combination with BSC) • IFN-α monotherapy (or in combination with BSC) • IL-2 monotherapy (or in combination with BSC) • Sunitinib monotherapy (or in combination with BSC) • Sorafenib monotherapy (or in combination with BSC) • Temsirolimus monotherapy (or in combination with BSC) • Bevacizumab in combination with IFN-α (and in combination with BSC)	• The included interventions are those which are either licensed for the first-line treatment of advanced/metastatic RCC or for which RCT data in this setting exist. • The review was limited to studies of these agents administered as monotherapy (or with the exception of bevacizumab in combination with IFN) as per their licensed indications or as per the anticipated license in the case of pazopanib.
	Comparator • Any of the included interventions • Placebo • Best supportive care*	• These comparators were chosen to enable both direct and indirect comparisons between the interventions of interest.
	Study design • Economic evaluations, including cost analyses, cost minimization analyses, cost-effectiveness analyses, cost utility analyses, utility studies.	• All economic evaluations should be considered.
	Language restrictions • English only	• The restriction would not limit results substantially due to data availability in English language.
	Publication timeframe • 1980 onwards for literature searches	• The restriction of date would not limit results substantially.
	Outcome of interest • Studies should report an outcome of interest. Outcomes of interest are: ○ Effectiveness and utilities ○ Resources ○ Costs ○ ICERs/ICs	• Studies which do not report outcomes of interest would not feature in any analyses or answer the review questions and were therefore were excluded.
Exclusion criteria	**No subgroup analysis** • No subgroup analysis for disease of interest • No subgroup analysis for advanced/metastatic disease • No subgroup analysis for treatment naïve patients	• Studies with no subgroup data for the disease, disease stage and line of treatment were not included, since these studies would introduce heterogeneity into the review.

*BSC definition: no active treatment/observation/a method of care that is not a focused treatment/treatments which clinicians consider to be "placebo-equivalent" including medroxyprogresterone and vinblastine. RCC = renal cell carcinoma

First pass of citations

Citations were first screened based on the abstract supplied with each citation. Each citation was screened by two independent reviewers, and any discrepancies between reviewers were reconciled by a third independent reviewer. Citations that did not match

the eligibility criteria were excluded at this 'first pass'; where unclear, citations were included. Duplicates of citations (due to overlap in the coverage of the databases) were also excluded in the first pass. Full-text copies of all references that could potentially meet the eligibility criteria were also ordered at this stage.

Second pass of citations

The eligibility criteria were applied to the full-text citations. Each full-text was screened by two independent reviewers, and any discrepancies between reviewers were reconciled by a third independent reviewer. Data presented in the studies still included after this stage were extracted to data extraction grids.

Extraction strategy

The final extraction grid is provided in Appendix C.2. Data from trials were extracted in parallel by two independent reviewers, with reconciliation of any differences by a third independent reviewer. All reviewers were qualified with either a Masters degree in pharmacy or an equivalent related discipline, and furthermore were fully trained in conducting systematic reviews with a minimum of 1.5 to 2 years of full-time experience of systematic review work within a health economics and outcomes research organisation.

Where more than one publication was identified describing a single trial, the data were compiled into a single entry in the data extraction table to avoid double counting of patients. Each publication was referenced in the table to recognize that more than one publication may have contributed to the entry.

Studies excluded during data each stage, along with rationale for exclusion are provided in a separate MS Excel document (Economic Excluded Studies).

3.2.4 Quality Assessment

Studies were assessed qualitatively using the Drummond and Philip's checklist (Drummond 1996; Philips 2004).

4 Overview of studies in the clinical systematic review

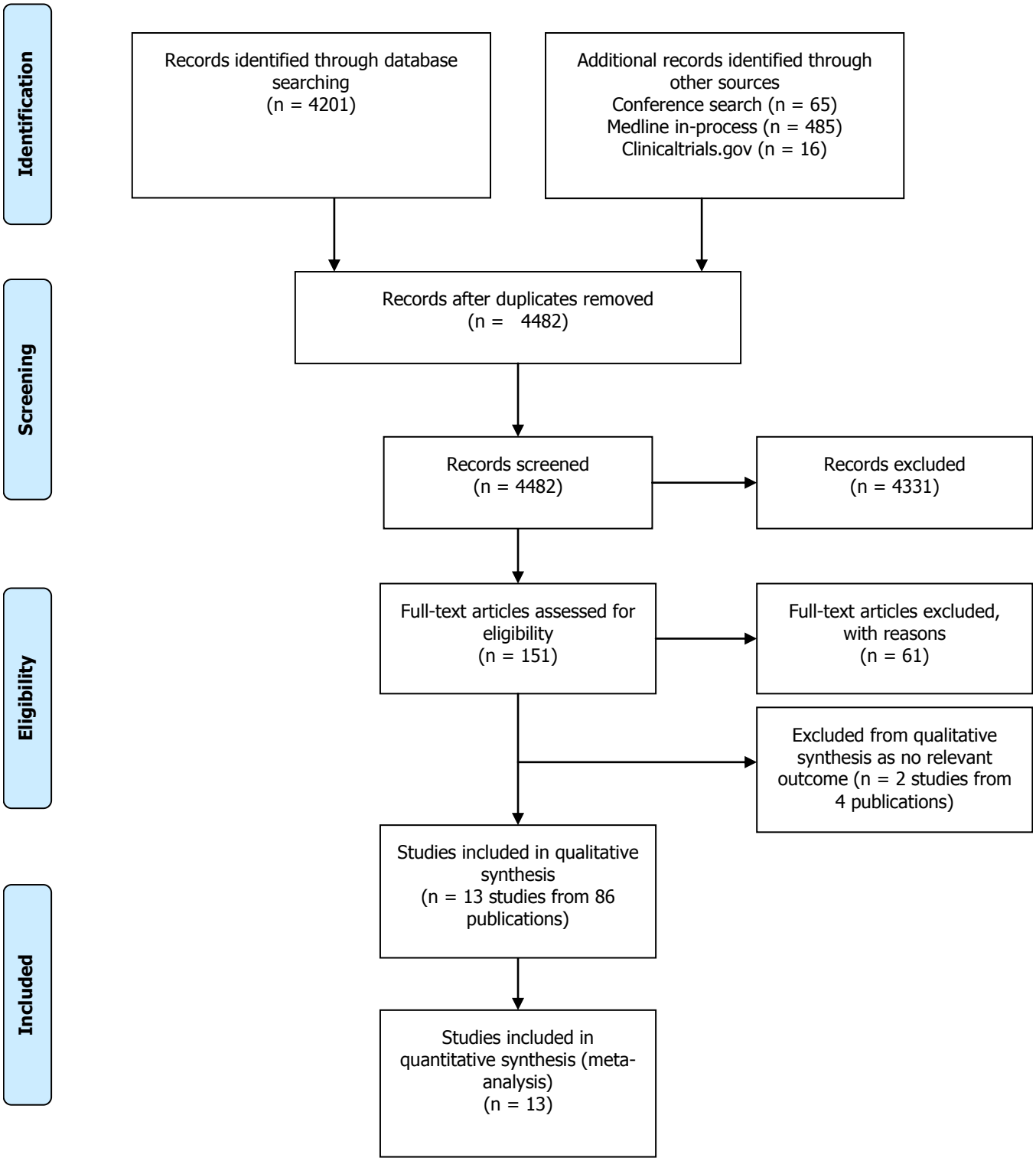
4.1 Trial Flow

The literature search yielded 4767 separate references. Due to the overlap of records across the databases searched, 285 references were found to be duplicates.

Following the first pass of the citations in the Heron SRDB, 151 potentially relevant references were identified. Full-text reports of these citations were obtained for more detailed evaluation.

Following detailed examination of the reports, 61 citations were excluded. Two studies, Soret 1996 and Prummer reported in four publications met the inclusion criteria but were excluded from the review since these did not report relevant outcomes (Soret 1996; Prummer 1994). Following linking of multiple publications per trial, 13 studies (86 publications) were extracted and contributed to the qualitative and quantitative analysis.

Figure 4: Trial Flow



4.2 Complete list of relevant randomised controlled trials (RCTs)

A total of 13 trials with relevant outcome data reported in 86 publications were identified as meeting the inclusion criteria for the review and were extracted. Additionally, one clinical study report (CSR) for pazopanib was also used for additional data to be extracted. This CSR reports data for a planned interim analysis for overall survival conducted with a clinical cut-off date of 23 May 2008. Final overall survival data are still awaited.

No studies which directly compared pazopanib with one of the comparators of interest (sorafenib, sunitinib, temsirolimus, Bevacizumab plus IFN, IFN or IL-2) were identified. One study comparing pazopanib with placebo was identified (VEG105192). In addition, studies comparing the comparators of interest to placebo or to IFN were however identified, permitting indirect comparison of interventions. The list of relevant RCTs included in this review and the treatment comparisons are provided in Table 10 and Figure 5. Of the studies which reported the phase, the majority were phase III trials while one phase II and one phase II/III trial was reported. Only two studies were double blind, with a third triple blinded. The remaining studies were either assessor blind, open-label or had an unclear level of blinding.

As expected with the number of phase III trials identified, the median number of participants enrolled in the included trials was large (435 patients, range 60 to 903). Additionally, most trials enrolled patients from more than one site (multicentre). Six of the included studies were conducted in more than one country (multicentre, international). In the included studies, targeted therapy was compared with immunotherapy in five trials while two trials compared targeted therapy with placebo. The remaining trials compared immunotherapy with BSC or another immunotherapy.

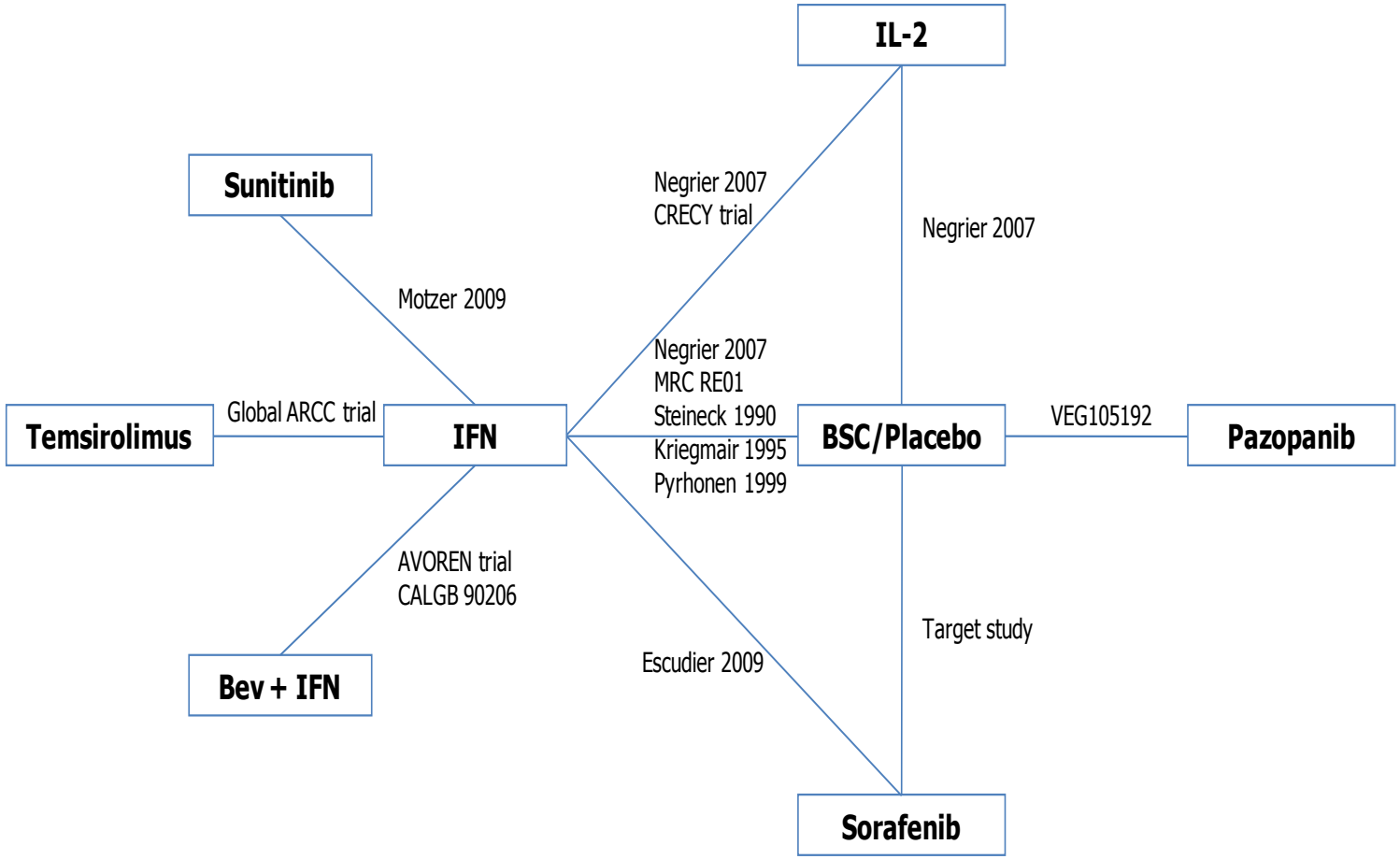
Table 10: List of relevant RCTs

Study	Year	Study type	N***	Intervention	Comparator	Patient population	Linked publications
Pazopanib							
#VEG105192 (GlaxoSmithKline 2008)	2009	R, DB, PC, MC-I, Phase III	233 treatment-naive (Total population = 435)	Pazopanib 800 mg od	Placebo	Locally advanced or metastatic clear cell/predominantly clear cell RCC, ECOG PS ≤ 1, Age ≥18 years	(Sternberg 2009b; Sternberg 2009a; Hawkins 2009b; Hawkins 2009a)
Sunitinib							
(Motzer 2009)	2009	R, AB, AC, MC-I, Phase III	750	Sunitinib 50 mg od	IFN 9 MU TIW	Metastatic RCC with a clear-cell histologic component, ECOG PS ≤ 1, Age ≥18 years	(Motzer 2008; Reddy 2006; Cella 2009; Patil 2009; Figlin 2008; Cella 2008a; Motzer 2007c; Eberhardt 2007; Cella 2008b; Motzer 2007b; Negrier 2008; Cella 2007a; Motzer 2007a; Motzer 2006a; Eberhardt 2006; Motzer 2006b; Castellano 2009)
Sorafenib							
(Escudier 2009c)	2009	R, OL, AC, MC, Phase II	189	Sorafenib 400 mg bid	IFN 9 MU TIW	Unresectable and/or metastatic, clear cell RCC, ECOG PS ≤ 1, Age ≥18 years	(Escudier 2006; Szczylik 2007)
#Target Study (Negrier 2009)	2009	R, TB, PC, MC-I, Phase III	161 treatment-naive (Total population = 903)	Sorafenib 400 mg bid	Placebo	Metastatic RCC, low or intermediate risk MSKCC score, ECOG PS 0 to 2, Age ≥18 years	(Autier 2008; Escudier 2009b; Eisen 2008; Bukowski 2009; Oudard 2009; Bukowski 2007b; Hutson 2009a; Eisen 2006; Escudier 2005; Escudier 2007a; Bellmunt 2007; Dhanda 2006; Jager 2005; Hutson 2009b; Bukowski 2007a)
Bevacizumab							
AVOREN trial (Escudier 2007c)	2007	R, DB, AC, MC-I, Phase III	649	Bevacizumab 10mg/kg q2wks plus IFN 9 MU TIW	Placebo plus IFN 9 MU TIW	Patients with clear-cell RCC and had undergone nephrectomy/partial nephrectomy, KPS of ≥70%, Age ≥18 years	(Melichar 2008; Melichar 2007; Escudier 2009a; Escudier 2008b; Melichar 2009; Bajetta 2008; Bellmunt 2009; Escudier 2007b; Bracarda 2007; Bracarda 2009; Ravaud 2008; Escudier 2008a)
CALGB 90206 (Rini 2008a)	2008	R, OL, AC, MC-I, Phase III	732	Bevacizumab 10mg/kg q2wks plus IFN 9 MU TIW	IFN 9 MU TIW	Metastatic RCC patients with clear cell histologic component, KPS of ≥70%, Age ≥18 years	(Rini 2004; Rini 2009; Rini 2008b)
Temsirolimus							
Global ARCC trial (Hudes 2007)[$]	2007	R, OL, AC, MC-I, Phase III	626	Temsirolimus 25 mg weekly[$]	IFN 18 MU TIW	Advanced RCC (stage IV or recurrent disease) and a KPS of ≥60%	(Dutcher 2009; Bellmunt 2008; Figlin 2009; Moore 2006; Alemao 2009;

Study	Year	Study type	N***	Intervention	Comparator	Patient population	Linked publications
							Mallick 2008; Parasuraman 2007; de Souza P. 2007; Dutcher 2007; Dutcher 2008; Logan 2008; Pendergrass 2009; Rajagopalan 2009; Yang 2009; de Souza 2008)
IFN, Interleukin-2							
(Negrier 2007)	2007	R, BU, AC, BSC, MC	492	IFN 9 MU TIW	Interleukin-2 9 MIU bid Medroxyprogesterone	Clearly progressive metastatic RCC of all histologic subtypes, >1 metastatic organ site and good performance status (KPS ≥80%) or 1 metastatic organ site with KPS 80%, Age ≥18 years	Negrier 2006
MRC RE01 (Ritchie 1999)	1999	R, BU, BSC, MC	350	IFN 10 MU TIW	Medroxyprogesterone	Histologically or cytologically confirmed metastatic RCC, WHO PS of 0 to 2	(Royston 2004; Royston 2008; Hancock 2000; Ritchie 1998)
(Steineck 1990)	1990	R, AB, BSC	60	IFN 10-20 MU/m2 TIW	Medroxyprogesterone	Locally recurrent or metastatic adenocarcinoma of kidney, Patients with previous irradiation of the disease or excision of metastases, Age 18 to 70 years	No links
(Kriegmair 1995)	1995	R, BU, BSC, Phase III	89	IFN 8 MU TIW plus vinblastine	Medroxyprogesterone	History of tumour nephrectomy and a histologically confirmed diagnosis of progressive RCC with bimensionally measurable tumour lesion and a WHO PS of at least grade 2	No links
(Pyrhonen 1999)	1999	R, BU, BSC, MC, Phase III	160	IFN 18 MU TIW plus vinblastine	Vinblastine	histologically or cytologically confirmed measurable or nonmeasurable but assessable advanced RCC, KPS >50% (ECOG status of 0 to 2), Age ≤75 years	(Hernberg 1997)
CRECY Trial (Negrier 1998)**	1998	R, AB, AC, MC, Phase II/III	425	Interkeukin-2 18 MU per m2 body surface area per day	IFN 18 MU TIW	Progressive metastatic RCC, ECOG PS<2, Age 18 to 65 years	(Negrier 1996; Lasset 1992)

*R = randomised, AB = assessor blind, AC = active controlled, BSC = best supportive care controlled, BU = blinding unclear, DB = double blind, ECOG = Eastern Cooperative Oncology Group, KPS = Karnofsky Performance Status, MC = multicentre, MC-I = multicentre-international, MSKCC = Memorial Sloan-Kettering Cancer Centre, MU = million units, od = once daily, OL = open label, PC = placebo controlled, TB = Triple blind, TIW = three times per week.
#subgroup analysis for treatment naïve patients; ***This is the number of treatment naïve patients in the study.
**This study also included an IFN-IL-2 combination treatment arm which was not extracted since it did not meet the inclusion criteria for intervention/comparator.
$ This study also included an IFN plus temsirolimus combination treatment arm which was not extracted since it did not meet the inclusion criteria for intervention/comparator.

Figure 5: Network diagram demonstrating treatment comparisons for all 13 included studies

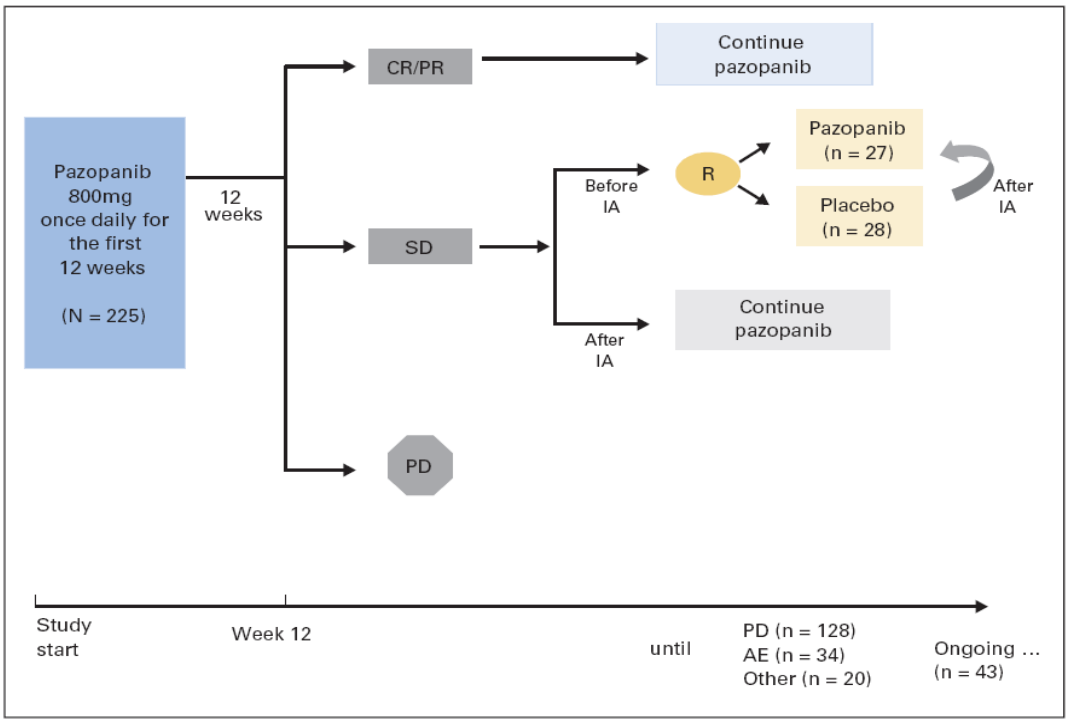


In addition to the one pazopanib RCT included in the review (see Table 10), a second pazopanib study was identified during the systematic review process. The second study was, however, excluded from the final list of included studies owing to the fact that (i) it was designed as a randomised discontinuation study but was later revised to a single-arm open-label study and (ii) lack of outcome data for the treatment naïve sub-group from the randomised phase of the study (Hutson 2007) VEG102616 study). A brief summary of the findings from this phase II study is provided below.

The study objective was to evaluate the safety and efficacy of pazopanib monotherapy (800 mg once daily) in comparison to placebo in patients with metastatic RCC who were either treatment-naïve or had failed prior cytokine or bevacizumab therapy. This phase II study was originally designed as a randomised discontinuation study but was revised to a single-arm open-label study on the recommendation of the data monitoring committee after a planned interim analysis gave an early indication of pazopanib's activity (response rate of 38% at 12 weeks in the first 60 patients). The study design is presented in Figure 6; all patients began the study on open-label pazopanib, and after 12 weeks of treatment patients with stable disease were to be entered into a randomised double-blind placebo-controlled phase. Patients with progressive disease (PD) discontinued the study and those with an CR or PR continued on open-label pazopanib. After the halt to randomisation all continuing patients were treated with open-label pazopanib. The primary end point was changed from PD rate at 16 weeks post-randomisation to the response rate.

The study enrolled 225 patients with metastatic RCC; 155 patients (69%) were treatment naïve, and 70 patients (31%) had received one prior cytokine- or bevacizumab-containing regimen. Approximately one third of enrolled patients had ECOG performance status score one and 41% belonged to the intermediate MSKCC risk category. Median time since diagnosis was 568 days.

Figure 6: Study flow chart



Taken from (Hutson 2010); IA = Interim analysis, CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, R = randomisation

Efficacy data for all 225 patients who received pazopanib or a combination of pazopanib

and placebo have been summarised together across the open-label and randomised phases of the study:

- Overall response rate (ORR) was 35% (95% CI: 28% to 41%)
- ORR was similar regardless of whether patients were treatment-naïve (34%, 95% CI: 26% to 41%) or had one previous line of therapy (37%, 95%CI: 26% to 49%).
- Median time to response (TTR) was 12 weeks and median duration of response (DOR) was 68 weeks.
- Median progression-free survival (PFS), adjusted for patients who were randomised to placebo was 52 weeks (95% CI: 44 to 60).
- Cytokine-pre-treated patients had a longer PFS (median 60.3 weeks) and treatment-naïve a shorter PFS (median 36.3 weeks) compared to the whole patient population, although the CIs for each estimate overlap, suggesting the difference may not be meaningful.
- ECOG performance status of 0 and time from diagnosis to treatment of more than 1 year were correlated with improved PFS.

Comparing median PFS for the randomised phase demonstrated pazopanib is associated with improved PFS versus placebo (52 and 27 weeks respectively; p=0.013).

Pazopanib was generally well tolerated in the trial. The most common AEs were diarrhoea, fatigue, and hair depigmentation. The most common laboratory abnormalities were elevated AST and ALT. The authors concluded that pazopanib demonstrated durable activity in patients with advanced RCC and was generally well tolerated in this population.

Following completion of the systematic review, an additional study was published and identified comparing the use of IFN and IL-2 in advanced metastatic RCC; MRC RE04/EORTC GU 30012 (Gore 2010). Although this trial did not meet the inclusion criteria for this review, a brief summary of the trial is provided below, since it is considered an important trial in this disease area, and is one of the largest conducted for IFN and IL-2.

The primary objective of MRC RE04 trial was to compare overall survival in patients receiving IFN alone or combination therapy with IFN, IL-2 and fluorouracil. The study was a randomised, open-label, multicentre-international trial conducted at 50 centres across eight countries. The study enrolled previously untreated patients with advanced/metastatic RCC with WHO performance status of 0 to 1. Approximately 90% of study participants had prior nephrectomy and the majority of study participants (approximately 59%) belonged to the intermediate MSKCC risk category.

Between April 2001 and August 2006, a total of 1006 patients were randomly allocated to receive IFN alone (502 patients) or combination therapy (504 patients) and followed prospectively. The median follow-up was 37.2 months (range, 24.8 to 52.3). The results of efficacy outcomes are as follow:

- There was no evidence of a difference between treatment groups for OS (HR 1.05, 95% CI: 0.90 to 1.21, p = 0.55).
- OS at 1 year was 67% in both treatment groups. Three year OS rate was 30% in patients receiving IFN alone compared to 26% in patients receiving combination therapy.
- Median PFS was 5.5 months in patients receiving IFN alone compared to 5.3 months in patients receiving combination therapy (HR = 1.02, 95% CI: 0.89 to 1.16, p = 0.81).
- Patients receiving combination therapy showed significantly higher ORR compared to patients receiving IFN monotherapy (23% vs. 16%, p = 0.0045).

Treatments were generally well tolerated and toxic effects were manageable. Grade 3/4 toxicity was reported more commonly in patients receiving combination therapy compared to IFN alone (53% vs. 36%, p<0.0001). Serious adverse events were reported in 23% patients receiving IFN alone compared to 26% patients receiving combination therapy. Authors concluded that combination therapy with IFN, IL-2 and fluorouracil provided no advantage compared to IFN alone in terms of OS or PFS.

4.3 Summary of trials in progress identified

The search of trials in progress resulted in 349 records. A total of 327 records did not meet inclusion criteria and were excluded. Reasons for exclusion were as follows:

- disease (144)
- study design (51)
- disease stage (1)
- line of therapy (6)
- intervention (92)
- control (39)
- copy/duplicate (1)
- trial status: completed/terminated (4)
- trial status: ongoing but not recruiting, results published (4)

Pazopanib is currently being investigated in a phase III randomised controlled trial versus sunitinib in the first-line treatment of advanced/metastatic RCC. Tabular summary of trials in progress is provided in Table 11.

Table 11: Summary of trials in progress

Trial ID	Title	Intervention	Comparator	Sponsors	N	Study design	Estimated completion date (month-year)	Outcome measure	Patient Population
NCT00720941 (VEG105884; COMPARZ)	Pazopanib versus sunitinib in the treatment of locally advanced and/or metastatic RCC	Pazopanib	Sunitinib	GSK	876	R, OL, MC-I, Phase III	Nov-10	PFS*, OS, Response rate, TTR, DOR, Safety, Health outcomes	Locally advanced and/or Metastatic RCC; Diagnosis of RCC with clear-cell component histology; no prior systemic therapy for advanced/metastatic RCC; KPS ≥70; Age ≥18 years
NCT00979966	Study in non-clear cell renal carcinoma (NCC-RCC) temsirolimus versus sunitinib	Temsirolimus	Sunitinib	Central European Society for Anticancer Drug Research	108	R, OL, MC, Phase II	Jul-11	TTP*, Response rate, PFS, OS, Safety	Locally advanced and/or Metastatic RCC; Diagnosis of RCC with non clear-cell histology; no prior systemic therapy for RCC; ECOG PS 0-2; Age ≥18 years
#NCT00619268	Combination of temsirolimus and bevacizumab in patient with metastatic RCC	Temsirolimus + Bevacizumab	Sunitinib Bevacizumab plus IFN	Centre Leon Berard	160	R, OL, MC, Phase II	Feb-12	PFSR*, Response rate, DOR, Toxicity, QoL, PFS, OS	Metastatic RCC; Diagnosis of RCC of all types except for papillary; no prior systemic therapy for metastatic RCC; ECOG PS ≤2; Age ≥18 years
#NCT00117637	BAY 43-9006 (sorafenib) versus IFN alpha-2a in patients with unresectable and/or metastatic RCC	Sorafenib	IFN	Bayer	189	R, OL, MC-I, Phase II	Apr-10	PFS*, DCR, OS, Response rate, PRO, Safety	Unresectable and/or metastatic RCC; predominantly clear cell RCC; prior surgical excision required; no prior systemic therapy for metastatic RCC; ECOG PS 0 pr 1; Age ≥18 years
#NCT00606866	MRI study of BAY 43-9006 in metastatic RCC	Sorafenib	Placebo	University of Chicago	57	RCO, DB, SC, Phase III	Jun-08	DCE-MRI*, Tumour shrinkage	Histologically or cytologically confirmed metastatic RCC; no prior anti-tumour kinase inhibitors or VEGF pathway inhibitors; ECOG PS 0-2; Age ≥18 years
#NCT00738530	A study of Avastin (bevacizumab) added to IFN alfa-2a (roferon) therapy in patients with	Bevacizumab plus IFN	IFN + Placebo	Hoffmann-La Roche	649	R, DB, MC-I, Phase III	July-10	OS*, PFS, TTP, TTF, Response rate, Safety	Metastatic clear-cell RCC; previous nephrectomy required; no prior systemic therapy for metastatic RCC;

Trial ID	Title	Intervention	Comparator	Sponsors	N	Study design	Estimated completion date (month-year)	Outcome measure	Patient Population
	metastatic RCC with nephrectomy								Age ≥18 years.

*primary outcome measure; N = estimated enrolment, Completion date = Estimated completion date, R = Randomised, RCO = Randomised cross over, OL = Open-label, DB = Double-blind, MC = Multi-centre, MC-I = Multi-centre international, OS = Overall survival, PFS = Progression free survival, TTP = Time to progression, TTF = Time to failure, DFS = Disease free survival, DOR = Duration of response, PRO = Patient reported outcomes, DCR = Disease control rate. QoL = quality of life
#trial status: ongoing but not recruiting participants, results not published

4.4 Methodology of relevant RCTs

4.4.1 Methods

A comparative summary of the methodology of the included RCTs is provided in Table 12. All studies were published as journal articles, with the exception of the pazopanib study, for which the CSR was provided (VEG105192).

All included studies reported the study location except the study by Escudier et al (Escudier 2009c). Additionally, all but one study (CALGB 90206 (Rini 2008a)) had at least one study site from a European country. Six of the included studies were international trials while six studies were regional trials enrolling patients from a single country (France, UK, Sweden, Finland or Germany).

All included studies were of similar study design (parallel RCT) though differed in terms of blinding, study setting and control group. The method of randomisation was largely unclear while it was adequate in three of the 13 included studies (AVOREN trial, CRECY trial and VEG105192). The TARGET study adopted a triple blind method while in three studies the assessor was blinded to treatment assignment (Steineck 1990, CRECY trial and Motzer 2009). Blinding status was unclear in four studies while three studies were described as an open-label trial (Escudier 2009c, Global ARCC trial and CALGB 90206).

Overall survival (OS) was the most common primary outcome measure (7 studies) in the included studies, followed by PFS (3 studies; Escudier 2009c, Motzer 2009 and VEG105192). Safety, quality of life and response rates were common secondary outcomes. Median follow-up duration varied across included studies with a minimum of 28.6 weeks in the TARGET study to a maximum of 242.67 weeks in the MRC RE01 study. The median follow-up in the pazopanib study at the time of the clinical cut-off was approximately 63 weeks; however the results of this trial are based on an interim analysis and the study is still ongoing.

Table 12: Comparative summary of methodology of the RCTs

Study	VEG105192	Motzer 2009	Escudier 2009	Target study	AVOREN trial	CALGB 90206	Global ARCC trial
Publication type	CSR	Journal articles					
Intervention	Pazopanib (N = 155)	Sunitinib (N = 375)	Sorafenib (N = 97)	Sorafenib (N = 77)	Bev + IFN (N = 327)	Bev + IFN (N = 369)	IFN (N = 207)
Comparator	Placebo (N = 78)	IFN (N = 375)	IFN (N = 92)	Placebo (N = 84)	IFN (N = 322)	IFN (N = 363)	Temsirolimus (N = 209)
Location	Argentina , Australia, Austria, Brazil, Chile, China, Czech Republic, Estonia, Hong Kong, India, Ireland, Italy, Korea, Latvia, Lithuania, New Zealand, Pakistan, Poland, Russia, Slovakia, Tunisia, Ukraine , UK	Australia, Brazil, Canada, France, Germany, Italy, Poland, Spain, UK, Russia, US	Unclear	Argentina, Australia, Belgium, Brazil, Canada, Chile, France, Germany, Hungary, Israel, Italy, The Netherlands, Poland, Russia, South Africa, Spain, Ukraine, UK, US	Australia, Belgium, Czech Republic, Finland, France, Germany, Hungary, Israel, Italy, The Netherlands, Norway, Poland, Russia, Singapore, Spain, Switzerland, Taiwan, UK	US and Canada	Argentina, Australia, Canada, Germany, Greece, Hungary, Czech Republic, Italy, Latvia, Lithuania, The Netherlands, Poland, Russia, Serbia and Montenegro, Slovakia, South Africa, Spain, Sweden, Taiwan, Turkey, Ukraine, UK, US
Design	R, DB, PC, MC-I, Phase III	R, AB, AC, MC-I, Phase III	R, OL, AC, MC, Phase II	R, TB, PC, MC-I, Phase III	R, DB, AC, MC-I, Phase III	R, OL, AC, MC-I, Phase III	R, OL, AC, MC-I, Phase III
Randomisation	Adequate	Unclear	Unclear	Unclear	Adequate	Unclear	Unclear
Blinding	Double blind, using matched placebo	Assessor-blind	Open-label	Triple-blind	Double blind	Open-label	Open-label
Primary outcomes*	PFS	PFS	PFS at 1 year	OS	OS	OS	OS
Secondary outcomes*	OS, DOR, Response rate, TTR, QoL, Safety, Withdrawals,	Response rate, OS, QoL, Safety, Withdrawals	QoL, Response rate, TTP, TTR, Safety, Withdrawals	PFS, Response rate, QoL, Safety, Withdrawals	PFS, Response rate, , TTR, TTP, Safety	PFS, Response rate, DOR, Safety, Withdrawals	PFS, Response rate, Clinical benefit rate, QoL, Safety, Withdrawals
Duration of follow-up	Median follow-up at clinical cut-off for interim analysis was 58.5 weeks (range, 3.9-97.93 weeks) for placebo group and 62.6 weeks (range, 1.73-106.17 weeks) for pazopanib group	Unclear	Unclear	Median 28.6 weeks	Median follow up was 99.23 weeks (22.9 months) in the in the Bev + IFN and 89.27 weeks (20.6 months) in the control group	Median follow up of censored patients was 200.02 weeks	Unclear
Timing of tumour assessment	Every 6 wks until week 24; every 8 wks thereafter	Day 28 of cycle 1-4; every 2 cycles thereafter	Every 8 weeks	Day 1 of cycle 2; every cycle thereafter and at 30 day follow-up visit	Every 8 weeks up to week 32; every 12 weeks thereafter	At baseline; every 12 weeks thereafter	Every 8 weeks

Study	Negrier 2007	CRECY Trial	MRC RE01	Steineck 1990	Pyrhonen 1999	Kriegmair 1995
Publication type	Journal articles					
Intervention	IFN (N = 122) IL-2 (N = 125)	IFN (N = 147)	IFN (N = 174)	IFN (N = 30)	IFN + BSC (N= 79)	IFN + BSC (N = 44)
Comparator	BSC (N = 123)	IL-2 (N = 138)	BSC (N = 176)	BSC (N = 30)	BSC (N = 81)	BSC (N = 45)
Location	France	France	UK	Sweden	Finland	Germany
Design	R, BU, AC, BSC, MC	R, AB, AC, MC, Phase II/III	R, BU, BSC, MC	R, AB, BSC	R, BU, BSC, MC, Phase III	R, BU, BSC, Phase III
Randomisation	Unclear	Adequate	Unclear	Unclear	Unclear	Unclear
Blinding	Unclear	Assessor-blind	Unclear	Assessor-blind	Unclear	Unclear
Primary outcomes*	OS	Response rate at 10 weeks	OS	Not identified	OS at 1 year and 5 years	Not identified
Secondary outcomes*	PFS, Response rate, QoL, Safety	OS, PFS, Safety	PFS, Response rate, Safety, QoL	Response rate, Safety	Response rate, DOR, TTP, Safety, Withdrawals	DOR, OS, Response rate, Safety, Withdrawals
Duration of follow-up	Median 126.53 weeks (range, 0 to 236.6 weeks)	Median 169 weeks	Median 242.67 weeks	Unclear	Unclear	Mean 39 weeks (IFN group; range, 4.33 to 104 weeks) and mean 27.3 weeks (BSC group; range, 4.33 to 95.33 weeks)
Timing of tumour assessment	12 weeks after start of treatment and between 24 and 26 in patients receiving further therapy	10 weeks after start of treatment and at week 25	At week 12 and 6 month	Every 4 weeks	Every 2 months	-

N = Number of patients randomised, R = Randomised, AB = Assessor blind, OL = Open-label, DB = Double-blind, TB = Triple blind, BU = Blinding unclear, MC = Multi-centre, MC-I = Multi-centre international, AC = Active controlled, PC = Placebo controlled, OS = Overall survival, PFS = Progression free survival, TTP = Time to progression, TTF = Time to failure, DFS = Disease free survival, DOR = Duration of response, QoL = Quality of life outcomes, TTR = Time to response

4.4.2 Participants

The detail of eligibility criteria is provided in Table 13. As per the review inclusion criteria, all studies enrolled adult patients (≥18 years) with advanced or metastatic RCC.

Both the Karnofsky Performance Scale (KPS) and Eastern Cooperative Oncology Group (ECOG)/WHO performance status are scales used to determine a patient's functional impairment and to determine how the disease is progressing. The Karnofsky score runs from 100 to 0, where 100 is "perfect" health and 0 is death (http://www.hospicepatients.org/karnofsky.html). While the ECOG performance status scale has six levels of activity: 0 – "Fully active" and 5 – "Dead". Patients with poor performance status are more prone to toxic reactions. Also, patients with a PS of 2 are reported to have a substantial incidence of grade 3 and 4 haematological and non-haematological toxicities (http://ecog.dfci.harvard.edu/general/perf_stat.html).

Patients in the included studies were generally required to have a Karnofsky Performance Status (KPS) score of at least 70-80% or an ECOG Performance Status score of ≤1, although this was not specified in all trials. Additionally, studies generally enrolled patients with a predominately clear cell histology (VEG105192, AVOREN trial and Global ARCC trial) whilst some studies restricted enrolment to only those with clear cell RCC (Motzer 2009, Escudier 2007 and CALBG 90206). Regarding, inclusion criteria, the temsirolimus study (Global ARCC trial, (Hudes 2007)) enrolled a patient population with a poor prognosis compared to the other studies; eligible patients were required to have at least three of six MSKCC prognostic risk factors. This difference in patient population for the temsirolimus study makes it less comparable to the other included studies, and this is clearly apparent when looking at baseline characteristics of enrolled patients (Table 14).

Patients with brain metastasis and those with cardiac disorders were generally excluded from the included studies. The majority of studies restricted entry of patients to those with normal hepatic, haematological and renal function. In addition, studies which enrolled patients with prior systemic treatment for metastatic RCC were excluded from this review unless a subgroup analysis of outcome data for treatment-naïve patients was presented. One study enrolled patients with prior systemic treatment; however, this study was included since the prior treatments in the cytokine-naïve sub-population were considered as equivalent to placebo or BSC (vinca alkaloids, pyrimidine analogues and progesterone) and an analysis of outcome data was identified for this sub-group (TARGET study, (Negrier 2009)). Additionally three studies enrolled small proportions of patients who were pre-treated with prior chemotherapy or hormonal therapy (CRECY trial (Negrier 1998), (Negrier 2007), (Pyrhonen 1999)).

Most of the trials allowed enrolment of patients with prior nephrectomy while three trials enrolled patients with prior nephrectomy only (AVOREN trial, CRECY trial and Kriegmair 1995). Studies which did not restrict enrolment to patients with only prior nephrectomy, nonetheless, enrolled a majority of patients who had undergone this surgery.

Table 13: Eligibility criteria in the RCTs

Study	Inclusion criteria	Exclusion criteria	Prior nephrectomy
Pazopanib			
VEG105192	• Adult patients with a diagnosis of clear cell or predominantly clear cell locally advanced (defined as disease not amenable to curative surgery or radiation therapy) or metastatic (Stage IV) RCC • Measurable disease presenting with at least one measurable lesion per RECIST • Cytokine pre-treated or treatment naïve disease • Adequate haematological, hepatic and renal function and • ECOG performance status 0 or 1 • At least 4 weeks had elapsed since the last surgery and 2 weeks had elapsed since radiotherapy or last systemic cytokine therapy at time of enrolment	• History of other malignancy • CNS metastasis • Malabsorption syndrome • Active peptic ulcer disease, inflammatory bowel disease, ulcerative colitis, or other gastrointestinal conditions with increased risk of perforation; abdominal fistula; gastrointestinal perforation, or intra-abdominal abscess within 4 weeks prior to beginning study treatment • History of HIV infection; uncontrolled infection • Cardiac angioplasty or stenting, or myocardial infarction, or unstable angina within the past 6 months • History of cerebrovascular accident or deep venous thrombosis (DVT) within the past 6 months • Poorly controlled hypertension; prolonged QTc interval	Approximately 84% of the patients had prior nephrectomy in the treatment-naïve subgroup.
Sunitinib			
Motzer 2009	• Patients aged ≥18 years with metastatic RCC with a clear-cell histologic component • Had not received previous treatment with systemic therapy. Presence of measurable disease • An ECOG performance status of 0 or 1 • Adequate hematologic, coagulation, hepatic, renal, and cardiac function	• Brain metastases • Uncontrolled hypertension • Clinically significant cardiovascular events or disease during the preceding 12 months	Prior nephrectomy was performed in 90% of the total randomised population.
Sorafenib			
Escudier 2009	• Patients aged ≥18 years with unresectable and/or metastatic, measurable and confirmed, predominantly clear cell RCC • No prior systemic therapy • Have ECOG PS ≤ 1 • Life expectancy ≥ 12 weeks • Complete surgical excision of primary PCC at initial diagnosis Adequate bone marrow, liver and renal function assessed 7 days before screening • Myocardial infarction ≥ 6 months before study entry were allowed and ß-blockers or digoxin were permitted	• Previous malignancy • Distinct in primary site/histology from that evaluated in this study • Complete renal failure that required dialysis • History of severe cardiac disease • Active, clinically serious bacterial or fungal infections • History of HIV, hepatitis B virus, or hepatitis C virus Symptomatic metastatic brain or meningeal tumour, seizure disorders that required medication • History of organ allograft, and substance abuse.	The study included those patients who had undergone complete surgical excision of primary RCC at initial diagnosis. Approximately 94.2% patients had undergone nephrectomy.
Target Study	• Patients aged ≥18 years with histologically confirmed metastatic RCC.	• Brain metastases • Previous exposure to VEGF pathway inhibitors	Approximately 93% of the patients had prior nephrectomy.

Study	Inclusion criteria	Exclusion criteria	Prior nephrectomy
	• Had undergone one prior systemic therapy (> 30 days and < 8 months from randomisation) • Low or intermediate risk MSKCC score • ECOG status 0 to 2, a life expectancy of ≥12 weeks • Adequate bone marrow, liver, pancreatic, and renal function • A prothrombin time or partial-thromboplastin time of < 1.5 times the upper limit of the normal range	• History of other malignancies • Organ allografts • Seizure disorders requiring medication • Clinically serious infections • Cardiac arrhythmias, symptomatic coronary artery disease, ischemia, or CHF	
Bevacizumab			
AVOREN trial	• Patients aged ≥18 years, with measurable or non-measurable tumour (RECIST criteria) • Had predominantly (>50%) clear-cell RCC • Had undergone nephrectomy/partial nephrectomy • Patients were required to have a KPS of ≥70% • Normal hepatic, hematopoietic and renal function	• Prior systemic treatment for metastatic RCC • Recent major surgical procedures • Evidence of brain metastases • Ongoing full-dose oral or parenteral anticoagulant or anti-platelet aggregation treatment • Uncontrolled hypertension on medication, clinically significant cardiovascular disease or chronic corticosteroid treatment	All patients had undergone nephrectomy or partial nephrectomy.
CALGB 90206	• Metastatic RCC patients • Aged ≥18 years with a clear-cell histologic component confirmed by local pathology review • No prior systemic therapy for RCC • KPS of ≥ 70% • Adequate bone marrow, hepatic, and renal function	• CNS metastases • New York Heart Association class II to IV heart failure Bleeding (haemoptysis, gastrointestinal bleeding) within 6 months • Blood pressure that could not be controlled to <160/90 mmHg with medication • History of venous thrombosis within 1 year, or arterial thrombosis (including cerebrovascular accident, unstable angina, myocardial infarction, or claudication with <1 block of exertion) within 6 months or who required ongoing therapeutic anticoagulation • Uncontrolled thyroid function • Requirement for systemic corticosteroids greater than physiologic replacement doses • Delayed healing of wounds, ulcers, or bone fractures	Approximately 85% of the patients had prior nephrectomy.
Temsirolimus			
Global ARCC trial	• Patients with histologically confirmed advanced RCC (stage IV or recurrent disease) • KPS of ≥60 • No previous systemic therapy • The tumour was to be measurable according to the RECIST Adequate bone marrow, renal, and hepatic functions were required • Patients had ≥3 of the following 6 prognostic risk factors: 1) <1 year from time of initial RCC diagnosis to randomization; 2)	• Patients with a history of brain metastases were eligible if their condition was neurologically stable • Did not require corticosteroids after surgical resection or radiotherapy.	Approximately 67% patients had undergone prior nephrectomy.

Study	Inclusion criteria	Exclusion criteria	Prior nephrectomy
	KPS of 60 or 70; 3) haemoglobin level less than the lower limit of the normal range; 4) corrected serum calcium level >10 mg/dL; 5) serum lactate dehydrogenase level >1.5 times the upper limit of the normal range; and 6) >1 metastatic organ site		
IFN, IL-2			
Negrier 2007	• Patients (≥18 years of age) with histologically confirmed, clearly progressive, metastatic RCC of all histologic subtypes • >1 metastatic organ and good performance status (KPS ≥80%) 1 metastatic organ with KPS 80% • Normal blood and liver functions with creatinine level ≤160 micromol/L	• Previous systemic treatment/ radiotherapy within 6 weeks of randomisation • Evidence of brain metastases • Uncontrolled cardiac dysfunction active infections • Current corticosteroid treatment • History of organ transplantation • Other cancer or seizure	Approximately 96% patients had undergone prior nephrectomy.
CRECY Trial	• Patients aged 18-65 years with histologically confirmed progressive metastatic RCC that could be measured in two dimensions • Patients had ECOG performance status of <2 • Normal blood cell counts • Normal bilirubin level, and creatinine levels below 1.7 mg per decilitre	• Brain metastases • Cardiac dysfunction • A contradiction to the use of vasopressor agents • Active infection • Previous treatment with interleukin-2 or IFN Chemotherapy or radiotherapy in the six weeks before enrolment • Current treatment with corticosteroids • Patients with a history of organ transplantation, other cancer, or seizure disorder	All patients had undergone nephrectomy or partial nephrectomy.
MRC RE01	• Patients with histologically or cytologically confirmed metastatic RCC • WHO performance status of 0 to 2	• Exclusion criteria were not reported in the study.	Patients with or without prior nephrectomy were included in the study; number of patients who underwent prior nephrectomy not stated.
Steineck 1990	• Patients with locally recurrent or metastatic adenocarcinoma of kidney • Aged between 18 years and 'a physiological age of 70' • With a life expectancy of > 12 weeks • Patients with previous irradiation of the disease or excision of metastases were included	• Patients with severe intercurrent disease • Any impaired function as judged by blood examinations	The initial protocol required a nephrectomy of the primary tumour but after the amendment of protocol this was not necessary; 3 patients in each group had not had the primary tumour excised prior to the outset of the trial.
Pyrhonen 1999	• Patients aged <75 years with histologically or cytologically confirmed measurable or non measurable but assessable advanced RCC • KPS >50% (ECOG status of 0 to 2) • Life expectancy >3 months • No abnormalities worse than mild (grade 1) in leukocyte, granulocyte, and platelet count, serum creatinine, and serum urea	• Brain metastases • Other malignancies • Serious concomitant illnesses • Radiotherapy involving more than 25% of the bone marrow reserve	71% patients in each group had undergone prior nephrectomy.
Kriegmair 1995	• Adult patients with a history of tumour nephrectomy	• Patients with fully resectable tumour lesions who underwent	All patients had undergone nephrectomy

Study	Inclusion criteria	Exclusion criteria	Prior nephrectomy
	• Histologically confirmed diagnosis of progressive RCC with dimensionally measurable tumour lesion • WHO performance status of at least grade 2	surgery and those with synchronous, bilateral tumour • Previous systemic treatment/radiotherapy • Other malignancies • Cardiovascular insufficiency (NYHA grade > 2) • Adequate hepatic, renal and blood function	or partial nephrectomy.

ECOG = Eastern Cooperative Oncology Group, KPS = Karnofsky performance status, RCC = renal cell carcinoma, RECIST = Response Evaluation Criteria In Solid Tumours criteria, WHO = World Health Organisation

Further detail regarding the patient characteristics of study participants at baseline are presented in Table 14 and Table 15.

All studies included men and women in a ratio of approximately 2:1, with a median age generally of around 60 years of age. Studies generally enrolled patients with an ECOG status of 0 or 1, although three IFN studies (Kriegmair 1995, MRCRE01 and Pyrhonen 1999) and the TARGET study, in which the sorafenib was compared with placebo, did include patients with an ECOG status of 2. Similarly, trials generally enrolled patients with a favourable or intermediate prognosis according to the MSKCC scoring system; however, in the study of temsirolimus vs. IFN (Hudes 2007) the majority (74%) of participants were of the MSKCC poor prognostic group. The predominant histological subtype was clear cell carcinoma in the included studies. Additionally, the majority of trial participants had undergone prior nephrectomy.

Table 14: Characteristics of participants in the RCTs across randomised groups in the pazopanib, sunitinib, sorafenib and bevacizumab trials

Study	**VEG105192**		**Motzer 2009**		**Escudier 2009**		**TARGET study**		**AVOREN trial**		**CALBG 90206**	
Intervention	**Pazopanib**	**Placebo**	**Sunitinib**	**IFN**	**Sorafenib**	**IFN**	**Sorafenib**	**Placebo**	**Bev + IFN**	**IFN**	**Bev + IFN**	**IFN**
N	155	78	375	375	97	92	77	84	327	322	369	363
Age (yrs)	59 (28-82)	62 (25-81)	62 (27-87)	59 (34-85)	62 (34-78)	62.5 (18-80)	60	60.5	61 (30-82)	60 (18-81)	61 (56-70)	62 (55-70)
Male (%)	68	74	71	72	67	56.5	63.6	69	68	73	73	66
Disease duration (yrs)	0.66	0.71										
ECOG performance status												
0	63 (40.6)	33 (42)	231 (61.6)	229 (61)	56 (58)	49 (53)	40 (52)	31 (37)			230 (62)	227 (62.5)
1	92 (59.4)	45 (58)	144 (38.4)	146 (39)	41 (42)	43 (47)	36 (47)	53 (63)			132 (36)	133 (36.6)
2							1 (1)	0			7 (2)	3 (1)
KPS												
100									144 (44)	124 (39)		
90									105 (32)	126 (39)		
80									58 (18)	50 (16)		
70									20 (6)	22 (7)		
MSKCC risk factors												
0 (favourable)	56 (36)	31 (40)	143 (38)	121 (32)	52 (53.6)	47 (51)	41 (53)	38 (45)	87 (27)	93 (29)	97 (26)	95 (26)
1-2 (intermediate)	87 (56)	40 (51)	209 (56)	212 (56.5)	44 (45.4)	44 (48)	36 (47)	46 (55)	183 (56)	180 (56)	234 (63)	231 (63.6)
≥ 3 (poor)	6 (4)	5 (6)	23 (6)	25 (6.7)	1 (1)	0 (0)			29 (9)	25 (8)	38 (10)	37 (10)
Histology												
Clear cell	135 (87)	69 (88.5)	375 (100)	375 (100)	97 (100)	92 (100)			278 (85)	283 (88)	369 (100)	363 (100)
Papillary												
Other												
Previous nephrectomy	130 (84)	65 (83)	340 (90.6)	335 (89)	95 (98)	83 (90)	70 (91)		327 (100)	322 (100)	312 (85)	308 (85)
Previous radiation therapy			53 (14)	54 (14.4)	22 (23)	12 (13)					35 (9.5)	38 (10.5)
No. metastases sites												
1	23 (15)	10 (13)	55 (14.7)	72 (19)	9 (9)	17 (18.5)						
2	46 (30)	25 (32)	106 (28)	112 (30)								
≥ 3	86 (55.5)	43 (55)	214 (57)	191 (51)								

*Dichotomous outcomes are reported as n (%) and continuous as median (range) unless otherwise specified.
ECOG = Eastern Cooperative Oncology Group, KPS = Karnofsky Performance Status, MSKCC = Memorial Sloan-Kettering Cancer Centre.

Table 15: Characteristics of participants in the RCTs across randomised groups in the temsirolimus, IFN and IL-2 trials

Study	**Global ARCC Trial**		**Negrier 2007**			**CRECY Trial**		**MRC RE01**		**Steineck 1990**		**Pyrhonen 1999**		**Kriegmair 1995**	
Intervention	**Temsirolimus**	**IFN**	**IL-2**	**IFN**	**BSC**	**IL-2**	**IFN**	**IFN**	**BSC**	**IFN**	**BSC**	**IFN + BSC**	**BSC**	**IFN**	**BSC**
N	209	207	125	122	123	138	147	167	168	30	30	79	81	44	45
Age (yrs)	58 (32-81)	60 (23-86)	61 (33-80)			56	55			63 (39-73)	62 (40-77)	60 (30-74)	62 (39-77)	62.4 (44-78)**	65.9 (47-79)**
Male (%)	66	71	75			69	73	72	65	70	80	65	63	63.64	68.89
Disease duration (yrs)										0.73	0.57	0.20	0.18		
ECOG performance status															
0			35			99 (72)	113 (77)	44 (26)	43 (25.6)			12 (15)	15 (18.5)		
1			65			35 (25)	30 (20)	83 (50)	80 (47.6)			53 (67)	49 (60.5)		
2						3 (2)	3 (2)	39 (24)	45 (27)			14 (18)	17 (21)	14 (32)	16 (35.6)
MSKCC risk factors															
0 (favourable)															
1-2 (intermediate)	64 (30.6)	50 (24)													
≥ 3 (poor)	145 (69)	157 (76)													
Histology															
Clear cell	169 (81)	170 (82)													
Papillary	25 (12)#	30 (14.6)#													
Other	40 (19)	37 (18)													
Previous nephrectomy	139 (66.5)	139 (67)	96			128 (93)	135 (92)	96 (57.5)	96 (57)			71 (90)	71 (88)	44 (100)	45 (100)
Previous radiation therapy			25			17 (12)	18 (12)					6 (7.6)	12 (15)	0	0
No. metastases sites															
1						31 (22.5)	41 (28)	28 (16.8)	26 (15.5)						
2						50 (36)	37 (25)								
≥ 3						56 (40.6)	69 (47)								
Prognostic factors: n (%)															
≥ 3	195 (93)	196 (95)													
< 3	14 (7)	11 (5)													

*Dichotomous outcomes are reported as n (%) and continuous as median (range) unless otherwise specified, **mean, # Based on N = 206 patients in IFN and temsirolimus groups
IFN = Interferon alpha, IL-2 = Interleukin 2, MSKCC = Memorial Sloan-Kettering Cancer Centre. KPS = Karnofsky Performance Status. ECOG = Eastern Cooperative Oncology Group

4.4.3 Outcomes

OS was the most commonly reported primary outcome measure (7 studies), followed by PFS (3 studies) in the included studies. Response rate was a primary outcome for one study though was a more commonly reported secondary outcome. Other commonly used secondary outcome measures were: TTP, TTR, DOR, safety, QoL and withdrawals. The list of primary and secondary outcomes reported in included studies is provided in Table 16.

Table 16: Primary and secondary outcomes of the RCTs

Outcome		VEG105192	Motzer 2009	Escudier 2009	Target Study	AVOREN trial	CALGB 90206	Global ARCC trial	Negrier 2007	MRC RE01	Pyrhonen 1999	CRECY Trial	Steineck 1990	Kriegmair 1995	Reliability/validity/current use in clinical practice (CHMP 2005)
Primary	OS				✓	✓	✓	✓	✓	✓	✓				OS is the gold standard outcome in late-stage oncology trials, including RCC, as it is a direct measure of clinical benefit that is unambiguously measured. The EMEA report that "acceptable primary endpoints include OS and PFS/DFS" and that "OS should normally be selected as the most appropriate primary endpoint." (CHMP 2005).
	PFS	✓	✓	✓											PFS is a reliable and valid surrogate outcome for determining the efficacy of treatments in RCC. The EMEA report that "PFS is an acceptable endpoint in situations where it is expected that further lines of treatment with effect on OS may importantly hamper the detection of a relevant treatment effect on OS." Since these trials are in first-line treatment, further lines may affect the true treatment effect, therefore this outcome could be considered valid in advanced/metastatic RCC (CHMP 2005). Sensitivity analyses are recommended to explore possible effects when events are detected between scheduled tumour assessments. This was conducted in the VEG 105192 Pazopanib study.
	Response											✓			The EMEA report that "without further justification, ORR is not an acceptable primary endpoint for confirmatory trials" (CHMP 2005).
Secondary	OS	✓	✓									✓		✓	The EMEA report that "Irrespective of the choice of primary endpoint OS/PFS, ORR and rate of tumour stabilisation for, e.g. 3 months should be reported" (CHMP 2005).
	PFS				✓	✓	✓	✓	✓	✓		✓			
	TTP			✓		✓					✓				The EMEA report that "alternative primary endpoints, such as TTP, TTF or EFS might uncommonly be appropriate." Therefore this could also be considered an appropriate secondary outcome.
	Response	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	The EMEA state that "irrespective of the choice of primary endpoint OS/PFS, ORR and rate of tumour stabilisation for, e.g. 3 months should be reported" (CHMP 2005). The report also states that "whenever possible, the definition of progression should follow established response evaluation criteria (e.g., RECIST)." This was used in a number of studies including

															the pazopanib trial.
	Safety	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	The EMEA report that "cumulative toxicity should always be investigated" (CHMP 2005).
	QoL	✓	✓	✓	✓			✓	✓	✓					The EMEA report that "in double-blind studies and especially in the palliative setting, HRQoL using generally accepted instruments might be valuable" (CHMP 2005). All trials used validated tools.
	Tolerability	✓	✓	✓	✓		✓	✓			✓			✓	The EMEA do not specifically recommend the reporting of these outcomes, however these may be useful in determining additional efficacy and safety of the interventions.
	TTR and DOR	✓		✓		✓	✓				✓			✓	

DOR = duration of response, QoL= Quality of life, OS= Overall survival, PFS= progression free survival, TTP= time to progression, TTR= time to response, RCC = renal cell carcinoma

4.4.4 Statistical analysis and definition of study groups

4.4.4.1 Statistical analysis

A summary of statistical analyses is provided in Table 17. Most of the included studies did not report their hypothesis clearly. All included studies, with the exception of two (Steineck 1990; Kriegmair 1995), reported sample size calculation. Most of the trials reported ITT analysis as primary analysis type and did not report the method employed for handling missing outcome data.

Table 17: Summary of statistical analyses in RCTs

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
Pazopanib				
VEG105192			The sample size calculation for OS was based on 90% power to detect a 50% improvement in median OS with pazopanib treatment compared with placebo. Given one interim analysis planned to occur after approximately 70% of the total events and flexible O'Brien-Fleming error spending functions for superiority and futility, this required accrual of 287 death events from approximately 350 enrolled subjects with a 2:1 randomisation. Upon amending the protocol to include the treatment-naïve subjects, shortly after the first subject was enrolled, the sample size was changed to 350 - 400 subjects to allow a minimum of 150 subjects to be enrolled for each of the treatment-naïve and cytokine pre-treated subgroups, and a minimum of 350 subjects to be enrolled for the entire study. This sample size allowed at least 90% power to detect an 80% improvement in median PFS by pazopanib treatment in both the overall study population and in each of the treatment-naïve and cytokine-pre-treated subgroups.	Patients were assessed on an ITT basis. The study was powered appropriately for the subgroup analysis of treatment naïve patients.
Sunitinib				
Motzer 2009			It was estimated from retrospective studies that 471 events (disease progression or death from any cause) would be required for 90% power to detect a clinically relevant increase in PFS from 4.7 to 6.2 months in patients treated with sunitinib, with the use of a two-sided, unstratified log-rank test with an overall significance level of 0.05. With a 1:1 randomisation of assignment to study groups, it was estimated that 690 patients are needed to enrol to observe 471 events. A total of 390 events were required for a two-sided, unstratified log-rank test with an overall two-sided significance level of p = 0.05 and 85% power to detect 35.7% improvement in overall survival.	The analysis included all patients randomly assigned to a study treatment group, according to an ITT basis. Additional exploratory analyses were performed to assess the treatment effect of sunitinib compared with IFN on overall survival, including censoring of patients at the date that the crossed over to sunitinib. The patients could exit the trial due to AEs, disease progression, consent withdrawal and other reasons.
Sorafenib				
Escudier 2009		Comparisons between treatment groups were	To achieve 85% power sufficient to detect a 66%	The efficacy analysis included all

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		performed using log-rank test and Cochran-Mantel-Haenszel test.	increase in PFS, data analysis was planned after approximately 140 PFS events were observed.	randomised patients (ITT).
Target Study		Treatment-related differences in response were evaluated by the Cochran–Mantel–Haenszel test.	Assuming a two-sided type I error of 0.04, the study had 90% power to detect a 33.3% difference in survival between the two groups after a total of 540 patients had died. Assuming that 3% of patients would be lost to follow-up, approximately 884 patients had to be randomly assigned to study groups.	
Bevacizumab				
AVOREN trial		Kaplan-Meier estimates were used to calculate OS and PFS. SAS version was used for statistical analysis.	The study was designed to have 80% power for the log rank test to detect an improvement in OS with an HR of 0.76, assuming an improvement of median survival from 13 to 17 months, at a two-sided alpha-level of 0.05. The planned sample size of 638 patients, with 445 deaths was required for the final analysis. One interim analysis was planned, after about 250 deaths had been observed. To ensure overall significance level at 5%, the interim analysis followed a sequential alpha spending function approach, using an O'Brien-Fleming boundary. With the planned interim analysis at 56% of the events, this approach resulted in a two-sided alpha-level of 0.0056 for the interim analysis and 0.0482 for the final analysis.	All patients who were randomised and exposed to study medication were included in the safety analyses.
CALGB 90206	The primary hypothesis was to investigate the clinical benefit of adding bevacizumab to IFN monotherapy.	The primary analysis on the overall survival end point was based on the stratified log-rank statistic. The primary analysis of the PFS end point was based on a two-sided stratified log-rank test comparing the two arms. In addition the Kaplan-Meier product-limit method was used to estimate the PFS time and DOR in the two arms. The threshold for significance for the PFS analysis was 0.05. The Chi-square test and Fisher's exact test were used to compare overall responses (ORRs) and AEs between the two treatment groups, respectively. All analyses were performed using SAS software.	The trial was designed with 86% power to detect 30% improvement in median survival in patients randomly assigned to Bevacizumab plus IFN compared with patients randomly assigned to IFN monotherapy, assuming a two-sided significance level of 0.05. The sample size calculations were based on the following assumptions were made: an annual accrual rate of 233 patients accrued over a 3-year enrolment period, 2-year follow-up period, and survival time following an exponential distribution.	The patients could exit the trial due to disease progression or death, toxicity, achieving CR, refusal to further treatment and other reasons.
Temsirolimus				
Global ARCC trial		The characteristics of the patients in each group were compared with the use of the chi-	The planned sample size of 200 patients per group was based on a power of 80% to detect a 40%	All patients who received any treatment were included in the

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		square test for categorical variables and the Kruskal-Wallis test for continuous variables. The proportion of patients with AEs in each group was analysed with the use of Fisher's exact test. Separate analyses were conducted for the comparison of the temsirolimus group with the combination-therapy group with the IFN group. All reported P values are two-sided and have not been adjusted for multiple testing.	improvement for each comparison with the use of a two-sided stratified log-rank test at an overall 2.5% level of significance	analysis of safety.
IFN, interleukin-2				
Negrier 2007	The hypotheses of survival at 2 years were 10%, 20% and 25% for patients assigned to MPA, IFN, and IL-2, respectively.	Survival estimates were calculated by using the Kaplan-Meier method. Differences in survival estimates between groups were assessed by the log-rank test with an assigned .025 level of significance for each of the main comparisons. Objective tumour response was assessed by descriptive statistical analysis; toxicity profiles tested by using the chi-square test; and quality of life assessed according to the EORTC QLQ-C30 scoring manual. Baseline scores for each quality-of-life domain and variations at Week 12 were calculated for each patient and compared between arms.	Assuming 80% power and .05 significance level, the planned sample size was 456 patients (114 in each of the four arms) and the number of expected deaths at the final analysis was 348. This sample size was also sufficient to detect a 15% difference between IL-2 and non-IL-2 groups.	
CRECY Trial			It was calculated that 138 patients per group were needed for a difference of at least 20% in overall survival to be detected, with alpha set at 5% and beta at 10%.	Analyses were performed on ITT population.
MRC RE01			The maximum sample size was set at 600 , and the triangular design guaranteed a power of 90% at 5% significance for detection of a difference in 2-year survival from 20% on MPA to 32% on IFN alpha (HR = 0.71). The trial was designed to be stopped after 150 – 250 deaths if no difference between treatments was apparent and after 200 – 300 deaths if the target improvement was shown with IFN.	Analysis was by intention to treat for primary efficacy outcome. Method for handling of missing data was not reported.
Steineck 1990	Primary hypothesis was not reported in the study.	Between treatment group comparison was made using Fisher's exact test and 2-sided p-value are reported.	Details of sample size and power calculations were not reported in the study.	Details of method used for account for missing data were not reported in the study.
Pyrhonen 1999	The hypothesis under	Overall survival, TTP and results were	The sample size calculation had 80% power to detect	All enrolled patients were assessable

Study	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
	consideration was that IFN alfa-2a added to a palliative regimen of vinblastine would prolong overall survival compared to vinblastine chemotherapy alone in patients with advanced RCC.	analysed according to Kaplan-Meier estimates and compared by use of the log-rank test, using life-table methods. The Cox proportional hazards model was applied to test for interactions between prognostic factors and survival within each treatment group and between prognostic factors and the effects of treatment on overall survival.	a difference between treatment groups in median survival of 12 versus 8 months, assuming recruitment of 160 patients at a rate of 40 patients per year and a follow-up period of 1 year.	for overall survival and TTP and results
Kriegmair 1995	The study did not state a primary hypothesis but focused on the survival benefit associated with the immunochemotherapy group.	Statistical analyses used were appropriate employing Chi-square analysis or Fisher's exact test for comparison of the distribution of tumour lesions and responders in both groups. Survival according to treatment schedule and response was estimated by the Kaplan-Meier method and was compared by the log rank test. p values <0.05 were considered significant for all tests.	Sample size calculation was not reported.	Type of analysis used for patients who withdrew was not reported. Treatment was discontinued in patients with tumour progression or with no change after a period of 3 months. Reason for discontinuation from the trial was reported as consent withdrawal.

CR = complete response, DOR = duration of response, EORTC QLQ-C30 = EORTC quality of life questionnaire C-30, HR = hazard ratio IFN = Interferon alpha, IL-2 = Interleukin-2, ITT = Intention to Treat Analysis, MPA = Medroxyprogesterone, QoL= Quality of life, OS= Overall survival, PFS= progression free survival, TTP= time to progression, TTR= time to response ,RCC = renal cell carcinoma.

4.4.5 Subgroup analyses

Two studies, namely VEG105192 and TARGET, presented subgroup analysis for treatment-naïve patients. In the case of VEG105192 study, the subgroup analysis was pre-specified and the study was designed with adequate power to detect a clinically meaningful increase in PFS for pazopanib compared with placebo in the treatment-naïve subgroup. The TARGET study did not specify the type of subgroup analysis (pre-specified or exploratory) and did not discuss about power of study with respect to subgroup analysis.

4.4.6 Participant flow

The summary of participant flow is provided in Table 18. All studies reported the number of patients randomised to each treatment arm. However, only two studies reported the number of patients screened for the study. Two studies reported subgroup analysis for treatment-naïve patients (TARGET study (Negrier 2009), VEG105192 (GlaxoSmithKline 2008)). All included studies with the exception of two, permitted cross-over of treatment in case of disease progression; detail regarding cross-over of treatment was not reported in three studies.

Table 18: Summary of participant flow

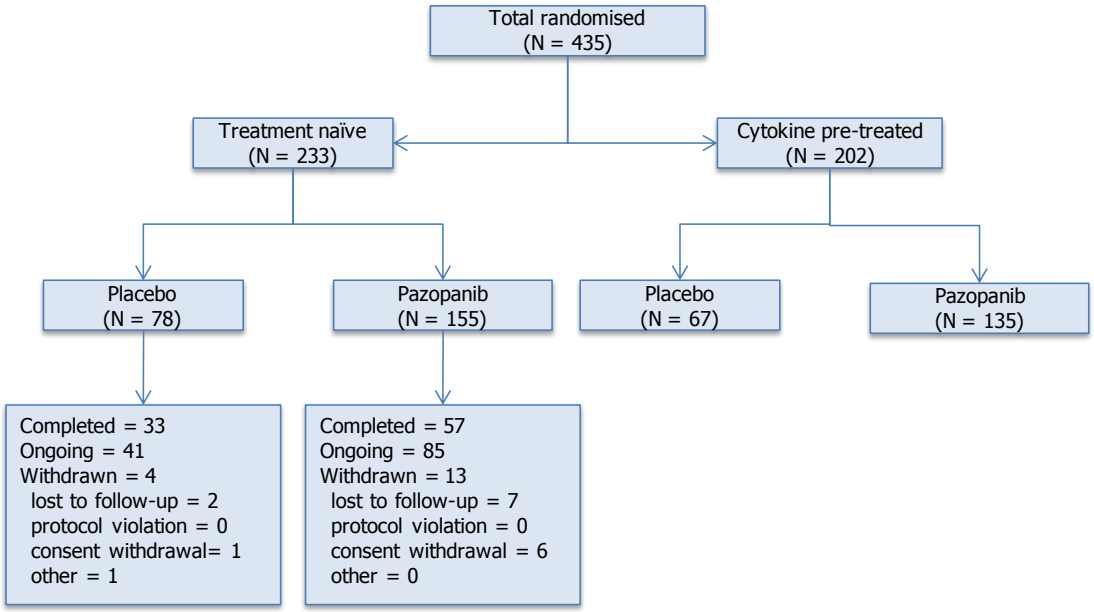
Study	Screened	Randomised	Completers	Cross over permitted?	
				YES/NO	Additional information
Pazopanib					
VEG105192		233**	90***	Yes	Patients who progressed were unblinded and if found to be on placebo were given an option to receive pazopanib through the open label extension study VEG107769. At time of interim OS analysis, 48% of all patients randomised to placebo had crossed over to receive pazopanib (40% of patients in the treatment-naive sub-group).
Sunitinib					
Motzer 2009		750	58	Yes	After the interim analysis had been performed and discussed with the data and safety monitoring committee, patients in the IFN group with progressive disease were allowed to cross over to the sunitinib group. About 7% of the patients (25 patients) in the IFN treatment arm crossed over to receive sunitinib while on study. Overall 33% of patients (117 patients) in the IFN arm received subsequent therapy with sunitinib (includes post study cancer treatment).
Sorafenib					
Escudier 2009		189		Yes	Patients who progressed on IFN were switched to sorafenib 400 mg twice daily within 14 days after IFN cessation.
Target Study		161*		Yes	Based on PFS benefit, crossover from placebo to sorafenib was permitted beginning in May 2005.
Bevacizumab					
AVOREN trial	821	649		Yes	After reviewing the final PFS interim

Study	Screened	Randomised	Completers	Cross over permitted?	
					OS results, the data and safety monitoring board recommended that patients in the control group who had not progressed should be crossed over to receive bevacizumab.
CALGB 90206	732	732		No	Cross over was not permitted, however patients with progressive disease received systemic anticancer therapy with other agents.
Temsirolimus					
Global ARCC trial		626		No	
IFN, interleukin-2					
Negrier 2007		492		Yes	Crossover between treatment groups occurred in 11.8% patients who were well distributed between arms
CRECY Trial		425		Yes	Patients who progressed could receive the other cytokine (cross over).
MRC RE01		350		Unclear	
Pyrhonen 1999		160	22	Unclear	
Kriegmair 1995		89		Unclear	
Steineck 1990		60		Yes	Fifteen patients were crossed over to IFN treatment after termination of medication with Medroxyprogesterone.

*Total population = 903; **Total population = 435; ***Ongoing (remaining in trial, follow-up ongoing) = 126. IFN = Interferon alpha.

The flow of study participants in pazopanib trial (VEG105192) is presented in Figure 7. Overall 17 patients withdrew from the study in the treatment-naive subgroup; 13 in pazopanib group and four in placebo group. Consent withdrawal and loss to follow-up were the major reasons behind premature termination of study.

Figure 7: Trial profile (VEG105192)



4.5 Critical appraisal of relevant RCTs

Summary of qualitative assessment is provided in Table 19 and

Figure 8. The complete quality assessment is provided in Appendix D. Methods used to generate random allocation sequence were reported in only three of the included studies and were judged as adequate; this included the AVOREN trial, CRECY trial and the VEG105192 pazopanib trial. Only five studies reported the method used for concealment of allocation sequence.

All of included studies reported comparable patient populations across interventions in the study. Evidence of selective reporting could not be determined for the majority of studies because of a lack of published protocol. Four studies had no evidence of selective reporting. However, since one or both OS and PFS were often measured and reported in each study this may not be an area of concern. All except one included study reported ITT analysis; however the method used to account for missing data was poorly reported.

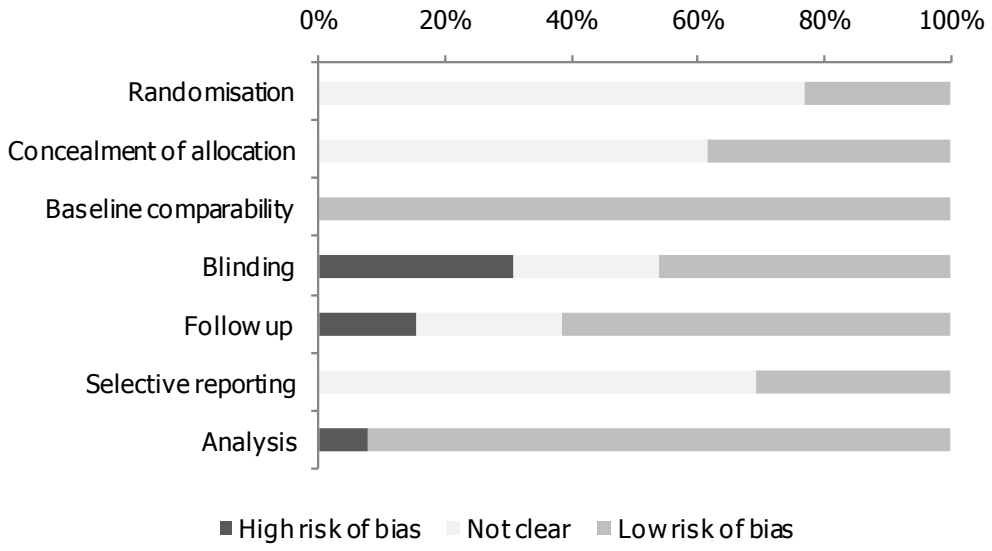
None of the studies were identified as being at a high risk of bias, so the validity of the results is not affected in each individual study. All studies were therefore included in the analyses of the review, where data availability permitted.

Table 19: Quality assessment results for RCTs

Study	Random-isation	Concealment grade	Baseline compar-ability	Blinding	Follow-up	Selective reporting	Analysis
Pazopanib							
VEG105192	Yes	Yes	Yes	Yes	No	No	Yes
Sunitinib							
Motzer 2009	Not clear	Not clear	Yes	Yes	No	Not clear	Yes
Sorafenib							
Escudier 2009	Not clear	Not clear	Yes	No	No	No	Yes
TARGET Study	Not clear	Not clear	Yes	Yes	No	Not clear	Yes
Bevacizumab							
AVOREN trial	Yes	Yes	Yes	Yes	Yes	Not clear	Yes
CALGB 90206	Not clear	Not clear	Yes	No	No	Not clear	Yes
Temsirolimus							
Global ARCC trial	Not clear	Not clear	Yes	No	No	No	Yes
IFN, IL-2							
Negrier 2007	Not clear	Yes	Yes	No	Not clear	No	Yes
CRECY Trial	Yes	Yes	Yes	Yes	Not clear	Not clear	Yes
MRC RE01	Not clear	Yes	Yes	Not clear	Not clear	Not clear	Yes
Steineck 1990	Not clear	Not clear	Yes	Yes	No	Not clear	Yes
Pyrhonen 1999	Not clear	Not clear	Yes	Not clear	No	Not clear	Yes
Kriegmair 1995	Not clear	Not clear	Yes	Not clear	Yes	Not clear	No

Randomisation; Was randomisation carried out appropriately? **Concealment grade;** Was the concealment of treatment allocation adequate? **Baseline comparability;** Were the groups similar at outset in terms of prognostic factors, for example, severity of disease? **Blinding;** Were the care providers, participants and outcome assessors blind to treatment allocation? **Follow-up;** Were there any unexpected imbalances in drop-outs between groups? **Selective reporting;** Is there any evidence to suggest that the authors measured more outcomes than they reported? **Analysis;** Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?

Figure 8: Risk of bias plot



5 Results of the studies in the clinical systematic review

5.1 Efficacy outcomes

All efficacy outcomes are discussed for the treatment-naïve population only.

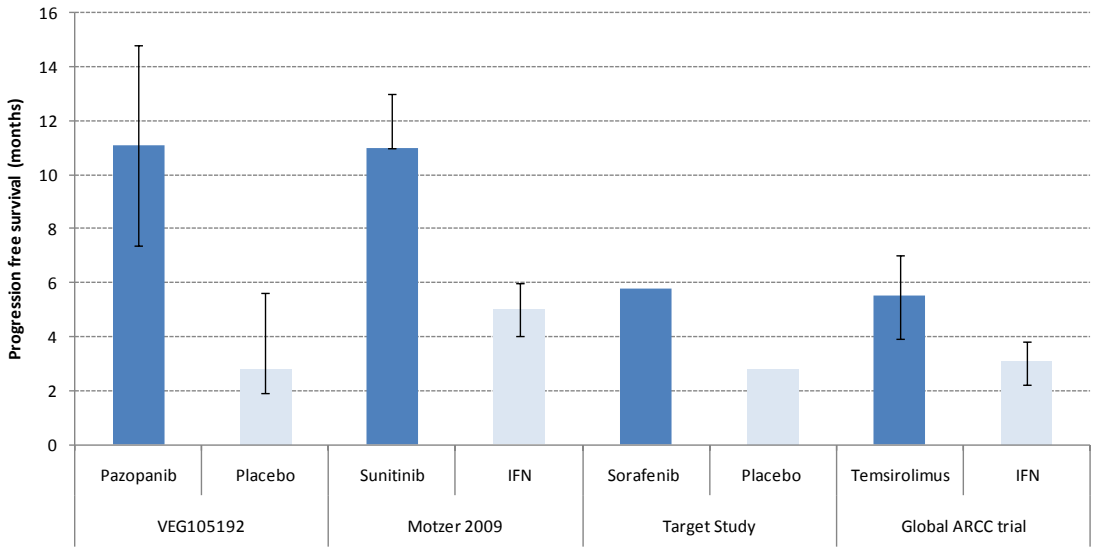
5.1.1 Progression free survival

Progression free survival (PFS) rate data were only available for the 1-year endpoint and in a limited number of studies, which prevented any meaningful trends from being observed. IFN and IL-2 demonstrated similar PFS rates of approximately 13% in one study (CRECY trial, (Negrier 1998)), while IFN demonstrated a much higher PFS rate at 1 year (30.4%) in a second study, which was much greater than that for sorafenib (11.5%) (Escudier 2009c).

Median PFS was much more widely reported in the included studies, and ranged from 3.4 months for IFN (Negrier 2007) to 11.1 months for pazopanib (VEG105192 trial) in active treatment groups, Table 20. In general, across all studies, median PFS was much greater in active treatment groups than comparators. The main exception was, however, in Escudier 2009, where the median PFS was almost identical in the sorafenib group and the IFN group. Results from the Negrier 2007 study also suggested that the difference between IFN, IL-2 and BSC with regard to median PFS, was minimal (0.4 months). Pazopanib, sunitinib and bevacizumab demonstrated the longest PFS in patients with advanced/metastatic RCC.

In addition, the PFS value for pazopanib was similar to that reported in the sunitinib study (Motzer 2009). The median PFS was much lower for temsirolimus than other comparator treatments; however it was greater than for the trial comparator, IFN. This is most likely attributable to a greater proportion of patients with poor prognosis.

Figure 9: Median PFS reported in the included studies (assessed by IRC)



IFN = Interferon alpha; 95% confidence intervals shown where reported in the publications.

In the VEG105192 trial, median PFS was significantly greater in the pazopanib group when compared to placebo with a hazard ratio (HR) of 0.36 (95% CI 0.24, 0.55; $p<0.001$). Data for PFS for pazopanib was taken from the sensitivity analysis which used scan date rather than clinical visit date. This data was used, since from personal communication with Motzer RJ, it was determined that PFS for sunitinib was also reported on the basis of scan date rather than clinical visit date.

Table 20: Summary of progression free survival

Study	Intervention	N (ITT)	1 year PFS rate % (n)	PFS in months, median (95% CI)	HR (95% CI)	Definition of PFS	Data Assessment
VEG105192	Pazopanib	155		11.1 (7.4, 14.8)	0.36 (0.24, 0.55), p<0.001#	Randomisation to the earliest date of documented disease progression or death due to any cause	Independent review committee (IRC)
	Placebo	78		2.8 (1.9, 5.6)			
Motzer 2009	Sunitinib	375		11 (11, 13)	0.539 (0.451, 0.643), p<0.001	Randomisation to first documented disease progression/death due to any cause.	Independent review committee (IRC)
	IFN	375		5 (4, 6)			
Escudier 2009	Sorafenib	97	11.5 (11)	5.7 (5, 7.4)		Not reported	Not reported
	IFN	92	30.4 (28)	5.6 (3.7, 7.4)	0.88 (0.61, 1.27), p = 0.504		
Target Study	Sorafenib	77		5.8	0.48 (0.32, 0.73)	Randomisation until the date of progression.	Independent review committee (IRC)
	Placebo	84		2.8			
AVOREN trial	Bev + IFN	327	43 (141)	10.2	0.63 (0.52, 0.75), p = 0.0001	Randomisation to first documented disease progression/death due to any cause.	Investigator
	IFN	322		5.4			
CALGB 90206	Bev + IFN	369		8.4	0.67 (0.57, 0.79), p<0.0001	Randomisation to first documented disease progression/death due to any cause.	Not reported
	IFN	363		4.9			
Global ARCC trial	Temsirolimus	209		5.5 (3.9, 7.0)** 3.8 (3.6, 5.2)***	0.74 (0.60, 0.90), p = 0.003***	Randomisation to disease progression or death, whichever occurred first	Independent review committee (IRC) and Investigator
	IFN	207		3.1 (2.2, 3.8)* 1.9 (1.9, 2.2)***			
Negrier 2007	IL-2	125		3.4 (2.9, 5.8)		Randomisation to first documented disease progression or death due to any cause.	Not reported
	IFN	122		3.4 (3, 5.6)			
	BSC	123		3 (2.9, 3.6)			
CRECY Trial	IL-2	138	15 (21)			Event free survival, defined as survival without disease progression.	Independent review committee (IRC)
	IFN	147	12 (18)				
MRC RE01##	IFN	174			0.66 (0.53, 0.82), p<0.001	Randomisation to first documented disease progression or death due to any cause.	Not reported
	BSC	176					

ITT N was used to calculate %. Bev = Bevacizumab, CI = confidence interval, ITT = Intention to treat analysis, PFS = progression free survival.
#this is a sensitivity analysis result for the pazopanib trial (based on scan date). *assessed by IRC (N = 153); **assessed by IRC (N = 192); ***investigator assessed
##updated result from Hancock 2000 is reported

5.1.2 Time to progression

Time to progression (TTP) was reported in only three (Escudier 2009, AVOREN trial and Pyrhonen 1999) out of the 13 included studies. TTP was defined as time from the date of randomisation to the time when progressive disease (PD) was observed.

In the AVOREN trial, TTP was significantly longer in the bevacizumab plus IFN group as compared to IFN group alone with a HR of 0.61 (p = 0.0001). In addition, Pyrhonen 1999 reported a significantly longer TTP in the IFN + BSC group compared to BSC alone. In comparison, however, Escudier 2009 reported similar TTP values in the sorafenib and IFN groups. No data for this outcome was reported in the pazopanib study since the primary outcome for this trial was PFS.

Table 21: Summary of TTP

Study	Intervention	N	TTP (median, months)	HR (95% CI)
Escudier 2009	Sorafenib	97	5.7	
	IFN	92	5.6	0.89 (0.61, 1.29)
AVOREN trial	Bev + IFN	327	10.2	0.61 (0.51, 0.73), p = 0.0001
	IFN	322	5.5	
Pyrhonen 1999*	IFN + BSC	81	3	
	BSC	81	2.08	

*TTP was calculated from the date of randomisation to the time when PD was observed. HR = hazard ratio
Bev = Bevacizumab, BSC= best supportive care, CI = Confidence interval, IFN= Interferon alpha

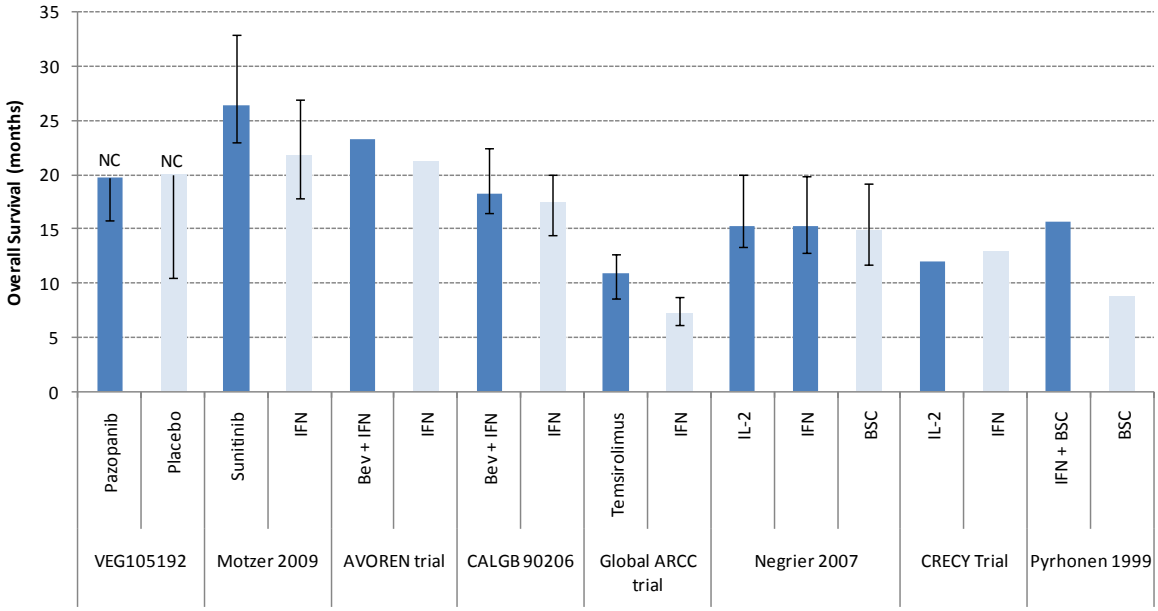
5.1.3 Overall Survival

Overall survival (OS) was the primary outcome in seven of the 13 included studies and was assessed from the date of randomisation to the date of death.

The rate of OS was available for the IFN versus BSC comparison only. In all studies for this comparison, the rate of OS at 1 year, 2 years and endpoint was greater for the IFN arm than BSC; although this difference was minimal in one study for the 2 year survival rate (Kriegmair 1995).

Median OS ranged from 9 months for IFN (MRC RE01) to 26.4 months for sunitinib (Motzer 2009) in active treatment groups, Figure 10. Median OS was greater in active groups than comparators across all studies except the CRECY trial, where median OS was slightly greater in IFN group as compared to the IL-2 group. In the VEG105192 trial, a similar median OS was reported in the pazopanib group and placebo group with a HR of 0.74 (95% CI 0.47, 1.15). It should be noted that at the time of data cut-off for this study, less than half the patients in the study had died and the maximum OS had not been reached in either of the groups.

Figure 10: Median overall survival reported in the included studies



NC = not calculable. The upper confidence interval for the VEG105192 trial was not calculable. Bev = Bevacizumab, BSC = best supportive care, IFN = Interferon alpha, IL-2 = Interleukin-2

The HR for median OS was reported in six of the included studies. The point estimates for HR ranged from 0.73 (Global ARCC trial, temsirolimus) to 0.86 (AVOREN, CALGB trial, bevacizumab) favouring the active drugs in all comparisons. Hazard ratios represented significant differences in OS between active and comparator groups in MRC RE01 and Global ARCC trial.

Table 22: Summary of overall survival

Study	Intervention	N (ITT)	OS rate % (n)			OS in months Median (95% CI)	HR (95% CI)
			1 year	2 year	endpoint		
VEG105192	Pazopanib	155	71.3 (111)			19.8 (15.8, not reached)	0.74 (0.47, 1.15)
	Placebo	78	59.6 (46)			20 (10.5, not reached)	
Motzer 2009	Sunitinib	375				26.4 (23, 32.9)	0.821 (0.673, 1.001), p = 0.051
	IFN	375				21.8 (17.9, 26.9)	
AVOREN trial	Bevacizumab plus IFN	327				23.3	0.86 (0.72, 1.04)
	IFN	322				21.3	
CALGB 90206	Bevacizumab plus IFN	369				18.3 (16.5, 22.5)	0.86 (0.73, 1.01), p = 0.069
	IFN	363				17.4 (14.4, 20)	
Global ARCC trial	Temsirolimus	209				10.9 (8.6, 12.7)	0.73 (0.58, 0.92), p = 0.008
	IFN	207				7.3 (6.1, 8.8)	
Negrier 2007	IL-2	125				15.3 (13.3, 20)	
	IFN	122				15.2 (12.8, 19.9)	
	BSC	123				14.9 (11.7, 19.2)	
CRECY Trial	IL-2	138				12	
	IFN	147				13	p = 0.55
MRC RE01#	IFN	174	43 (75)	22 (38)		9	0.75 (0.53, 0.82), p = 0.013
	BSC	176	32 (56)	13 (23)		6+	
Pyrhonen 1999	IFN + BSC	79	55.7 (44)		4.1 (3)	15.6	
	BSC	81	38.3 (31)		0 (0)	8.72	
Kriegmair 1995	IFN + BSC	44		18 (8)			
	BSC	45		16 (7)			

ITT N was used to calculate %. HR = hazard ratio. CI = confidence interval, ITT = intention to treat, OS = overall survival
#updated results from Hancock 2000 presented.

5.1.4 Response rate

Response rate was reported in all 13 studies, commonly as a secondary outcome (in 12 out of the 13). Response rate was assessed using Response Evaluation Criteria In Solid Tumours (RECIST, (Therasse 2000)) and World Health Organisation (WHO) criteria in seven and five studies, respectively while the criteria of assessment was not reported for one study (MRC RE01). The comparison of the two criteria are summarised in Table 23.

Table 23: WHO and RECIST criteria for tumour response (Park 2003)

	WHO	**RECIST**
Measurability	Measurable, bidimensional	Measurable, unidimensional: Conventional method ≥ 20 mm; Spiral CT ≥ 10 mm; Target versus non-target lesion
	Non-measurable/evaluable	Non-measurable
Objective response		
Complete response	Disappearance of all known lesion(s); confirmed at 4 weeks	Disappearance of all known lesion(s); confirmed at 4 weeks
Partial Response	At least 50% decrease; confirmed at 4 weeks	At least 30% decrease; confirmed at 4 weeks
Stable disease	Neither PR nor PD criteria met	Neither PR nor PD criteria met
Progressive disease	25% increase; no CR, PR or SD documented before increased disease, or new lesion(s)	20% increase; no CR, PR, or SD documented before increased disease, or new lesion(s)

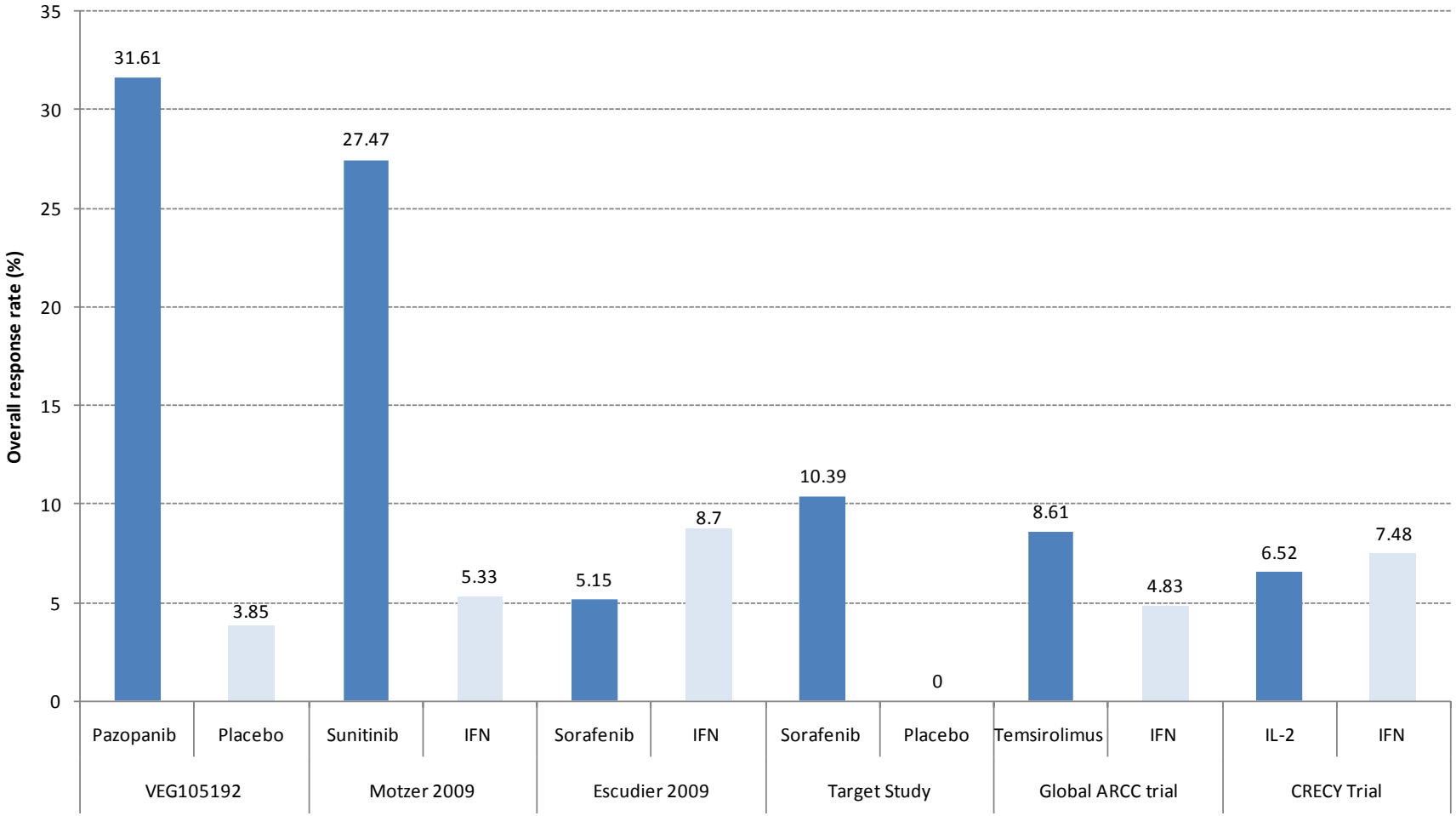
Response rate was assessed by independent review committee (IRC) in six studies and by investigator in four studies (3 studies report both) while it was unclear in six studies. Overall response rates (ORR) rarely exceeded 10% for the IFN, IL-2 studies, Figure 11. The proportion of patients achieving an ORR as assessed by the IRC was greater in patients receiving pazopanib in the VEG105192 study compared with those receiving sunitinib in the sunitinib trial (Motzer 2009). Temsirolimus demonstrated an increased ORR compared to IFN. However, the percentage was much lower than for the other treatments of interest which is most likely attributable to the poorer prognosis patient population in this trial.

Table 24: Summary of response rate

Study	Intervention	N (ITT)	Response rate % (n)								Criteria for response assessment
			Assessed by Independent review committee				Assessed by Investigator/unclear				
			ORR	CR	PR	SD	ORR	CR	PR	SD	
VEG105192	Pazopanib	155	31.61 (49)	0 (0)	31.61 (49)	36.13 (56)	38.71 (60)	1.29 (2)	37.42 (58)	35.48 (55)	RECIST criteria
	Placebo	78	3.85 (3)	0 (0)	3.85 (3)	39.74 (31)	6.41 (5)	0 (0)	6.41 (5)	43.59 (34)	
Motzer 2009*	Sunitinib	375	27.47 (103)	0 (0)	27.47 (103)	42.67 (160)	46.93 (176)	2.93 (11)	44.00 (165)	40.00 (150)	RECIST criteria
	IFN	375	5.33 (20)	0 (0)	5.33 (20)	42.67 (160)	12.27 (46)	1.07 (4)	11.20 (42)	53.87 (202)	
Escudier 200	Sorafenib	97	5.15 (5)	0 (0)	5.15 (5)	74.23 (72)					RECIST criteria
	IFN	92	8.70 (8)	1.09 (1)	7.61 (7)	55.43 (51)					
Target Study	Sorafenib	77	10.39 (8)	1.30 (1)	9.09 (7)	75.32 (58)					RECIST criteria
	Placebo	84	0 (0)	0 (0)	0 (0)	55.95 (47)					
AVOREN trial	Bev + IFN	327					29.36 (96)	1.22 (4)	28.13 (92)	43.12 (141)	RECIST criteria
	IFN	322					11.49 (37)	1.86 (6)	9.63 (31)	44.72 (144)	
CALGB 90206	Bev + IFN	369	Of the total 639 patients with measurable disease, overall response rate was higher in patients treated with Bevacizumab plus IFN (25.5%, 95% CI: 20.9%, 30.6%) than for those treated with IFN monotherapy (13.1%, 95% CI: 9.5%, 17.3%; p<0.0001).								RECIST criteria
	IFN	363									
Global ARCC trial	Temsirolimus	209	8.61 (18)				8.61 (18)				RECIST criteria
	IFN	207	4.83 (10)				7.73 (16)				
Negrier 2007	IFN	122					8.20 (10)	2.46 (3)	5.74 (7)	18.85 (23)	WHO criteria
	IL-2	125					4.00 (5)	0 (0)	4.00 (5)	20.80 (26)	
	BSC	123					1.63 (2)	0.81 (1)	0.81 (1)	14.63 (18)	
CRECY Trial	IL-2	138	6.52 (9)	1.45 (2)	5.07 (7)	21.74 (30)					WHO criteria
	IFN	147	7.48 (11)	0 (0)	7.48 (11)	31.29 (46)					
MRC RE01#	IFN	174					6.32 (11)	1.15 (2)	5.17 (9)	12.64 (22)	Not reported
	BSC	176					2.27 (4)	0 (0)	2.27 (4)	8.52 (15)	
Steineck 1990	IFN	30					6.67 (2)	3.33 (1)	3.33 (1)	13.33 (4)	WHO criteria
	BSC	30					3.33 (1)	3.33 (1)	0 (0)	10.00 (3)	
Pyrhonen 1999	IFN + BSC	79					16.46 (13)	8.86 (7)	7.59 (6)	39.24 (31)	WHO criteria
	BSC	81					2.47 (2)	1.23 (1)	1.23 (1)	43.21 (35)	
Kriegmair 1995	IFN + BSC	44					20.45 (9)	9.09 (4)	11.36 (5)	25.00 (11)	WHO criteria
	BSC	45					0 (0)	0 (0)	0 (0)		

ITT N was used to calculate %; Highlighted cells (yellow) means data assessment is unclear (not reported if investigator assessed or assessed by independent review committee); * Number analysed = 662 (335 in sunitinib group and 327 in IFN group); Bev = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha, ITT = intention to treat; ORR = overall response rate, CR = complete response, PR = partial response, SD = stable disease; #extracted from Ritchie 1999 as updated results were not reported in Hancock 2000

Figure 11: Summary of overall response rates (assessed by IRC) reported in the included studies



IFN = interferon alpha, IL-2 = interleukin-2

5.1.5 Time to, and duration of, response

The time to response (TTR) was poorly reported in the included studies; only three studies presented data for this outcome (VEG105192, AVOREN trial, Escudier 2009), Table 25. Pazopanib study defined time to response as the time from randomization until the first documented evidence of CR or PR (whichever status was recorded first). All studies showed that the active treatment, pazopanib, bevacizumab plus IFN and sorafenib were associated with shorter TTR than comparator treatment.

Table 25: Summary of TTR

Study	Intervention	N	TTR in weeks Median (range)
VEG105192	Pazopanib	49	11.6 (95% CI: 6.4, 12.3)
	Placebo	3	23.6 (95% CI: 18.1, 24.1)
Escudier 2009	Sorafenib	5	7.8 (7.37, 16.03)
	IFN	8	23.4 (16.03, 47.67)
AVOREN trial	Bev + IFN	298	9.53 (8.67, 60.67)
	IFN	276	16.03 (4.33, 43.33)

IFN = interferon alpha

Data for the duration of response (DOR) were also poorly reported (Table 26). DOR was calculated according to WHO criteria in one study (Pyrhonen 1999). In both AVOREN and CALGB trials, a longer DOR was reported with combination of IFN and bevacizumab compared to the control groups. Sunitinib showed longer DOR compared to IFN however the difference was not statistically significant. The median DOR for pazopanib was observed to be higher than that reported for sunitinib.

Table 26: Summary of DOR

Study	Intervention	N	DOR (months)	Definition	Comments
VEG105192	Pazopanib	49	13.55 (95% CI: 10.36, 15.25)	Time from first documented evidence of CR or PR until the first documentation of disease progression or death due to any cause, whichever was first.	
	Placebo	3	Not reached (95% CI: 8.70, not reached)		
Motzer 2009	Sunitinib	165	12 (95% CI: 10, 14)	Not Reported	
	IFN	43	10 (95% CI: 8, 17)		
AVOREN trial	Bev + IFN	298	13.5 (range, 1.8 to 20.3)	Not Reported	
	IFN	276	11.1 (range, 3.7 to 19.5)		
CALGB 90206	Bev + IFN		11.9 (95% CI: 8.3, 14.8), p = 0.977	Not Reported	N unclear
	IFN		8.7 (95% CI: 5.6, 11.4)		N unclear
Pyrhonen 1999	IFN + BSC	13		DOR was calculated using standard WHO criteria.	The median DOR was 27 weeks or 6.23 months for the seven patients who achieved CR (range, 12 to 281 weeks) and 24 weeks or 5.54 months for the six patients who achieved PR (range, 18 to 63 weeks).
	BSC	11			The median DOR was 100+ weeks or 23.08+ months for a patient who achieved CRs and 24 weeks or 5.54 months for a patient who achieved PRs.
Kriegmair	IFN + BSC	9		Not Reported	Mean duration of remission

Study	Intervention	N	DOR (months)	Definition	Comments
1995					for complete responders and partial responders was 10.8 months (range: 4 to 16) and 11.6 months (range: 7 to 20) respectively.
	BSC				

Bev = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha.

5.2 Health related quality of life

Health related quality of life outcomes were reported in only three studies (Motzer 2009, Escudier 2009 and Global ARCC trial) in addition to the pazopanib study (VEG105192). Various HRQoL tools were used in these studies and are summarised in table 26.

Table 27: Summary of HRQoL tools used

HRQoL Tool	Studies using this tool	Validation paper
EQ-5D	VEG105192; Motzer 2009; Global ARCC trial	http://www.euroqol.org/home.html
EQ-VAS	Motzer 2009	
EORTC-QLQ C-30	VEG105192	(Aaronson 1993)
FACT-Kidney Symptom Index–Disease-related Symptom (FKSI-DRS Index)	Motzer 2009	(Cella 2007b)
FACT-Kidney Symptom Index - 15 item scale (FKSI-15 Index)	Motzer 2009; Escudier 2009	(Cella 2006)
Functional assessment of cancer therapy – general scale (FACT-G)	Motzer 2009	(Lee 2004)

5.2.1 EQ-5D

EQ-5D Index score is a reliable and valid tool for the assessment of health-related quality of life. The EQ-5D Index score ranges from –0.594 to 1.000, with scores of 1, 0, or less than 0 denoting that the corresponding health state is valued by the population as equivalent to full health, death, or worse than death, respectively. EQ-5D score was reported in two of the included studies.

In the pazopanib study (VEG105192), results from a mixed-model repeated measures (MMRM) analysis for change from baseline consistently showed no statistical difference in EQ-5D score between pazopanib and placebo arms at each assessment time point (Table 28).

Table 28: Summary of EQ-5D (VEG105192)

Change from baseline to	N	Mean (SE)		Difference (95% CI) vs. placebo
		Pazopanib	Placebo	
Week 6	202	-0.037 (0.02)	-0.027 (0.03)	-0.010 (-0.081, 0.061), p = 0.784
Week 12	166	-0.044 (0.02)	-0.034 (0.03)	-0.010 (-0.080, 0.061), p = 0.789
Week 18	136	-0.017 (0.02)	-0.020 (0.03)	0.003 (-0.067, 0.073), p = 0.930
Week 24	116	-0.023 (0.02)	-0.015 (0.04)	-0.008 (-0.094, 0.079), p = 0.861
Week 48	60	0.020 (0.02)	-0.006 (0.04)	0.026 (-0.059, 0.111), p = 0.548

In a study comparing sunitinib with IFN, the overall post-baseline mean treatment difference was estimated to be 0.364 points in favour of sunitinib (95% CI: 0.0109 to 0.0620, p = 0.0364), Table 29. Similar results were reported for trial comparing temsirolimus with IFN, where a significant difference of 0.098 points was observed in favour of temsirolimus (95% CI: 0.036 to 0.162, p = 0.0022).

Table 29: Summary of EQ-5D

Study	Intervention	N	Baseline	Endpoint
Motzer 2009 (reported in Cella 2008)	IFN	356	0.76 ± 0.23	0.73 (N = 319)
	Sunitinib	373	0.76 ± 0.23	0.762 (N = 349)
Global ARCC trial	IFN	155	0.62 ± 0.24	0.49 (SE = 0.031)
	Temsirolimus	115		0.59 (SE = 0.026)

IFN = Interferon alpha

A significant benefit of sunitinib over IFN was observed for QoL in the Cella 2008 study, while no significant difference was observed for pazopanib compared to placebo. However, the pazopanib study compared the mean difference from baseline in EQ-5D scores for pazopanib and placebo, while Cella 2008 compared only endpoint scores between treatments. The authors of the sunitinib study also report that "results predominately reflected between-group differences rather than within-group improvement from baseline", therefore may not have demonstrated significant results given the same analysis as that conducted in the pazopanib trial. In addition, although the authors report that the concern is mitigated, patients in the sunitinib trial were aware of their treatment allocation, while patients in the pazopanib trial were blinded to their assigned therapy.

5.2.2 EQ-VAS

The EQ-VAS is a 100-point VAS (0 = worst imaginable health state; 100 = best imaginable health state) that expresses the patient's self-perceived value for his/her health state. EQ-VAS score was reported in only one study comparing sunitinib with IFN. The overall mean treatment difference was estimated to be 4.74 points in favour of sunitinib (95% CI: 2.60 to 6.87, p<0.0001), Table 30.

Table 30: Summary of EQ-VAS

Study	Intervention	N (BL)	Baseline	Endpoint	Comments
Motzer 2009 (reported in Cella 2008)	IFN	356	71.43 ± 19.51	68.7 (N = 319)	Least square mean reported; estimated from mixed-effects model, and the average post-baseline score were computed at approx. week 17.
	Sunitinib	373	73.8 ± 18.5	73.4 (N = 349)	

IFN = Interferon alpha

5.2.3 EORTC-QLQ-C30

The EORTC QLQ-C30 is a 30-item self-reporting questionnaire developed to assess the quality of life of cancer patients. The QLQ-C30 incorporates nine multi-item scales: five functional scales (physical, role, cognitive, emotional, and social); three symptom scales (fatigue, pain, and nausea and vomiting); and a global health and quality-of-life scale.

The EORTC-QLQ-C30 score was reported in only one study comparing pazopanib with placebo (VEG105192). Results from a MMRM analysis for change from baseline consistently showed no statistical difference between pazopanib and placebo arms at each assessment time point in global health status/HRQOL (Table 31).

Table 31: Summary of EORTC-QLQ-C30 (VEG105192)

Change from baseline to	N	Mean (SE)		Difference (95% CI) vs. placebo
		Pazopanib	Placebo	
Week 6	192	-5.03 (1.567)	-2.75 (2.342)	-2.28 (-7.859, 3.299), p = 0.421
Week 12	162	-3.91 (1.573)	-3.58 (2.534)	-0.33 (-6.231, 5.573), p = 0.913
Week 18	134	-4.25 (1.664)	-1.31 (2.804)	-2.95 (-9.401, 3.510), p = 0.369
Week 24	114	-2.25 (1.690)	-1.12 (2.957)	-1.12 (-7.870, 5.622), p = 0.742
Week 48	60	-0.79 (2.051)	-1.59 (3.593)	0.80 (-7.404, 9.014), p = 0.845

5.2.4 FACT-Kidney Symptom Index–Disease-related Symptom (FKSI-DRS Index)

The FACT-Kidney Symptom Index–Disease-related Symptom subscale (FKSI-DRS) is a subscale of the validated FACT-Kidney Symptom Index–15 item scale (FKSI-15) that contains nine items measuring symptoms predominantly related to kidney cancer. The FKSI-DRS score ranges from 0 (all most severe symptoms) to 36 (no symptoms). Only one study reported FKSI-DRS index (Motzer 2009). In this study patients in sunitinib group reported higher (more favourable) FKSI-DRS scores than those in the IFN group. Overall mean difference in scores was 1.98 points (95% CI: 1.46 to 2.51) favouring sunitinib group (p<0.0001).

Table 32: Summary of FKSI-DRS Index

Study	Intervention	N	Mean ± SD score	
			Baseline	Endpoint
Motzer 2009 (reported in Cella 2008)	IFN	356	29.55 ± 5.03	27.4 (N = 319)
	Sunitinib	373	29.74 ± 5.24	29.4 (N = 349)

IFN = interferon alpha.

5.2.5 FACT-Kidney Symptom Index - 15 item scale (FKSI-15 Index)

The FKSI-15, introduced in 2006, is a validated symptom index for kidney cancer patients containing 15 questions, each scored on a 5-point scale (0 = not at all; 4 = very much) (Cella 2006). The FKSI-15 score ranges from 0 (most severe symptoms and concerns) to 60 (no symptoms or concerns). The secondary publication of Motzer 2009 (Cella 2008), reported the HRQoL outcomes. In this study, sunitinib was compared with IFN. Patients in sunitinib group reported higher (more favourable) FKSI-15 scores than those in the IFN group (Table 33). Overall mean difference in scores was 3.27 points (95% CI: 2.36 to 4.18) favouring sunitinib group (p<0.0001). Similar results were reported for trial comparing sorafenib with IFN, where a clinically significant difference of 5.9 points was observed in favour of sorafenib (p = 0.015).

Table 33: Summary of FKSI-15 Index

Study	Intervention	N	Mean ± SD score	
			Baseline	Endpoint
Escudier 2009	IFN	92		34.6
	Sorafenib	97		40.5
Motzer 2009 (reported in Cella 2008)	IFN	356	46.1 ± 8.7	42.1 (N = 319)
	Sunitinib	373	46.45 ± 8.46	45.3 (N = 349)

IFN = interferon alpha, SD = standard deviation

5.2.6 Functional assessment of cancer therapy – general scale (FACT-G)

FACT-G, a reliable and valid scale (Lee 2004), measures the impact of treatment on general cancer related Health Related Quality of Life (HRQoL) and functioning. In a study comparing sunitinib with IFN, patients in sunitinib group reported higher (more favourable) FACT-G scores than those in the IFN group. Overall mean difference in scores was estimated to be 5.58 points (95% CI: 3.91 to 7.24) favouring sunitinib group (p<0.0001).

Table 34: Summary of FACT-G

Study	Intervention	N	Mean ± SD score	
			Baseline	Endpoint
Motzer 2009 (reported in Cella 2008)	IFN	356	81.25 ± 16.04	76.8 (N = 319)
	Sunitinib	373	82.3 ± 15.2	82.3 (N = 349)

FACT-G = functional assessment of cancer therapy – general guide, IFN = Interferon alpha

5.3 AEs

5.3.1 Any AE

A comparative summary of the percentage of patients who experienced any grade AE and any grade 3/4 AE in each treatment arm is provided in Table 35. As expected, owing to the nature of the interventions, AEs (all grades) were commonly reported, though were slightly less frequent with pazopanib (VEG105192) compared to the combination regime of bevacizumab plus IFN (AVOREN trial) and sorafenib (Escudier 2009c). Overall AEs (all grades) were not reported in the study conducted by Motzer 2009.

The frequency of grade 3/4 AEs was low with pazopanib in comparison to bevacizumab plus IFN, sorafenib, and temsirolimus. The frequency of grade 3/4 AEs ranged from 14.6% (Kriegmair 1995) to 79% (CALGB trial, (Rini 2008a)) in active intervention groups and from 0% (Kriegmair 1995) to 78% (Global ARCC trial, (Hudes 2007)) in comparator groups. A lower rate of serious AEs was reported with pazopanib as compared to bevacizumab plus IFN (AVOREN trial, (Escudier 2009c)).

Table 35: AEs reported by randomised patients (overall)

Study	Intervention	N	Any AE irrespective of grade % (n)	Grade 3/4 AEs % (n)	Comments
Any AE					
VEG105192	Pazopanib	155	90.97 (141)	42.58 (66)	
	Placebo	78	74.36 (58)	19.23 (15)	Grade 5 toxicity was reported in two patients.
Escudier 2009	Sorafenib	97	94.85 (92)	41.24 (40)	
	IFN	90	88.89 (80)	35.56 (32)	
Target Study	Sorafenib	77	72.73 (56)	19.48 (15)	
	Placebo	83	48.19 (40)	3.61 (3)	
AVOREN trial	Bev + IFN	337	97.33 (328)	60.24 (203)	Grade 3 or worse AE
	IFN	304	94.41 (287)	45.07 (137)	Grade 3 or worse AE
CALGB 90206	Bev + IFN	366		78.96 (289)	G-III 66.39 (243); G-IV 12.59 (46)
	IFN	349		61.03 (213)	G-III 56.45 (197); G-IV 4.58 (16)
Global ARCC trial	Temsirolimus	208		67 (139)	Calculated from 67% of 208; range of values possible (139 to 140)
	IFN	200		78 (156)	Calculated from 78% of 200; range of values possible (155 to 156)
Negrier 2007	IFN	122		40.16 (49)	
	IL-2	124		62.1 (77)	
	BSC	121		9.92 (12)	
Kriegmair 1995	IFN	41		14.63 (6)	Grade 3 fever
	BSC	35		0 (0)	
Pyrhonen 1999	IFN + BSC	79		18.99 (15)	Grade IV toxicity reported
	BSC	81		2.47 (2)	Grade IV toxicity reported
Any serious AE					
VEG105192	Pazopanib	155	21.29 (33)		
	Placebo	78	16.67 (13)		
Escudier 2009	Sorafenib	97	16.49 (16)		
	IFN	90	14.44 (13)		
AVOREN trial	Bev + IFN	337	29.08 (98)		
	IFN	304	16.45 (50)		
Any treatment related AE					
Escudier 2009	Sorafenib	97	94.85 (92)		
	IFN	90	88.89 (80)		
MRC RE01	IFN	174	50.57 (88)		Assessed at week 12
	BSC	176	11.36 (20)		Assessed at week 12

Evaluable N was used to calculate %; G-III = Grade III, G-IV = Grade IV; **Dark green** (0) represents 0% patients reported AE, **Light green** (1) represents 1-25% patients with AE, **Yellow** (2) represents 26-50% patients with AE, **Orange** (3) represents 51-75% patients with AE, **Red** (4) represents 76-100% patients with AE.
AE = adverse event, Bev = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha, ITT = intention to treat

5.3.2 Specific AEs

A summary of specific AEs experienced at any grade by patients randomised to each intervention is provided in Table 36. AEs are grouped by class, and demonstrate that the interventions of interest are generally associated with a relatively low (<25%) risk of musculoskeletal, nervous system, respiratory and psychiatric disorders.

AEs in the blood and lymphatic system class of disorders, such as anaemia, are however more common, particularly for sunitinib and IFN. Additionally, sunitinib appears to be associated with an increased risk of AEs compared to pazopanib, bevacizumab, sorafenib and temsirolimus, since the risk of certain classes of AEs exceed 50%. In fact, the incidence of anaemia, leucopenia and neutropenia exceeded 75% in patients treated with sunitinib. On qualitative analysis, it was observed that higher proportion of patients had increased AST and ALT levels after the treatment with pazopanib as compared with sunitinib.

There were fewer apparent differences between treatments for the more severe grades of AEs (grade 3 or 4), Table 37. However, pazopanib was associated with few grade 3 or 4 AEs. For several AEs, none of the patients treated with pazopanib experienced a more severe grade, whereas a small number of patients experienced these AEs whilst receiving other treatments.

Table 36: Specific AEs experienced by randomised patients (across all grades)

AEs by class		VEG105192		Motzer 2009		Escudier 2009		Target Study		AVOREN trial		Global ARCC trial		Steineck 1990	
		PAZO	PLAC	SUN	IFN	SORA	IFN	SORA	PLAC	BEV + IFN	IFN	TEM	IFN	IFN	BSC
	N	155	78	375	360	97	90	77	83	337	304	208	200	30	30
GI disorders	Abdominal pain	12.3	1.3	11	3	8.2	4.4					21	17		
	Diarrhoea	47.1	6.4	61	15	54.6	12.2	27.3	10.8	20.5	15.5	27	20	3.3	0
	Dyspepsia	3.9	1.3	31	5										
	Vomiting	21.9	5.1	31	12	13.4	14.4	5.2	1.2			19	28		10
	Nausea	25.8	10.3	52	35	18.6	27.8	13	14.5			37	41	3.3	0
	Mucositis/stomatitis	3.2	0	30	4			5.2	0			20	4		
General disorders	Asthenia	16.8	7.7	20	19					32.3	27.6	51	64		
	Fatigue	18.7	12.8	54	52	43.3	43.3	24.7	13.3	32.6	27.3			66.7	10
	Fever	7.1	5.1	8	35	3.1	32.2			45.1	42.8	24	50		
Skin and subcutaneous tissues disorders	Alopecia	9	0	12	9	41.2	5.6	29.9	3.6					3.3	0
	Hair colour change	38.7	1.3	20	1										
	Hand-foot syndrome	1.9	0	29	3	59.8	4.4	22.1	3.6						
	Rash	7.7	3.8	24	8	41.2	8.9	31.2	12			47	6	10	3.3
	Skin discolouration	5.8	0	27	1										
Investigations	ALT increased	25.2	2.6	51	40										
	AST increased	20	2.6	56	38							8	14		
	Total bilirubin increased	1.9	1.3	20	2									3.3	0 (0)
Vascular disorder	Hypertension	39.4	9	30	4	22.7	5.6	6.5	0	26.1	9.2			0	6.7
Metabolism and nutrition disorders	Anorexia	25.2	10.3	34	28	29.9	30	10.4	8.4	35.9	30.3	32	44	20	3.3
	Hyperglycaemia	2.6	0									26	11		
	Hypophosphataemia	0.7	0	31	24										
Musculoskeletal and connective tissue disorders	Arthralgia	6.5	2.6	11	14										
	Flank pain	0	1.3												
Nervous system disorders	Altered taste	8.4	1.3	46	15										
	Headache	13.5	5.1	14	16					23.4	16.1	15	15		
Respiratory, thoracic and medistinal disorders	Epistaxis	1.3	0	18	2									3.3	0
Infections and infestations	Infection	22.6	17.9									27	14		
	Flu-like symptoms	2.6	2.6			4.1	22.2			24.3	25.3			100	3.3
Blood and lymphatic system disorders	Anaemia	3.2	7.7	79	70			1.3	0	9.8	13.5	45	42	40	30
	Leucopenia	3.2	0	78	57							6	17	46.7	0
	Lymphocytopenia	1.3	0	68	69										

AEs by class		VEG105192		Motzer 2009		Escudier 2009		Target Study		AVOREN trial		Global ARCC trial		Steineck 1990	
		PAZO	PLAC	SUN	IFN	SORA	IFN	SORA	PLAC	BEV + IFN	IFN	TEM	IFN	IFN	BSC
	Neutropaenia	5.2	0	77	50			0	0	7.1	6.6	7	12		
	Thrombocytopaenia	7.7	1.3	68	26			0	0	6.2	3.9	14	8	6.7	0
Psychiatric disorders	Depression	2.6	1.3			0	14.4			12.2	10.2				
Cardiac disorders	Congestive heart failure	0.6	0							0.3	0.3				
Endocrine disorders	Hypothyroidism	5.2	0	14	2										

Evaluable N was used to calculate %. No AE data for specific AE's for CALGB 90206, Negrier 2007, CRECY trial, MRC RE01, Kriegmair 1995 and Pyrhonen 1999.

Dark green represents 0% patients reported AE, **Light green** represents 1-25% patients with AE, **Yellow** represents 26-50% patients with AE, **Orange** represents 51-75% patients with AE, **Red** represents 76-100% patients with AE.

PAZO = pazopanib. SUN = Sunitinib. SORA = sorafenib. BEV = bevacizumab. TEM = temsirolimus. IFN = Interferon. Bev + IFN = Bevacizumab plus interferon. BSC = Best supportive care.

Table 37: Specific Grade 3/4 AEs experienced by randomised patients

Study	VEG105192		Motzer 2009[1]		Escudier 2009[2]		Target Study		AVOREN trial[2]		CALGB 90206[1]		Global ARCC trial		Negrier 2007			CRECY Trial[3]		MRC RE01[4]		Kriegmair 1995		Pyrhonen 1999[5]	
	PAZO	PLAC	SUN	IFN	SORA	IFN	SORA	PLAC	BEV + IFN	IFN	BEV+ IFN	IFN	TEM	IFN	IFN	IL-2	BSC	IFN	IL-2	IFN	BSC	IFN + BSC	BSC	IFN + BSC	BSC
N	155	78	375	360	97	90	77	83	337	304	366	349	208	200	122	124	121	147	138	51	49	41	35	79	81
Abdominal pain	2.6	0	2/0		3.1	1.1							4	2											
Diarrhoea	3.2	0	9/0	1/0			1.3		2.1	1.0			1	2	0	4	0	0.7	27.5						
Dyspepsia	0	1.3	2/0	<1/0																7.8	32.7				
Vomiting	3.2	0	4/0	1/0	2.1	1.1							2	2	0.8	7.3	0.8								
Nausea	1.3	0	5/0	1/0		3.3					7.1/0	4.6/0	2	4	0.8	5.6	0			11.8	12.2				
Mucositis/stomatitis	0	0	1/0	<1/0			3.9	0					1	0	1.6	0	0								
Asthenia	0	0	7/<1	4/0					10.1	6.6			11	26											
Fatigue	1.9	5.1	11/0		5.2	10	1.3		11.9	8.2	34.7/1.9	28.1/1.7								58.8	53.1				
Fever	0	0	1/0	<1/0					2.4	0.7			1	4	3.3	11.3	0	5.4	42.8			14.63	0		
Alopecia	0	0																				0	0		
Hair colour change	0.7	0																							
Hand-foot syndrome	0	0	9/0	1/0	11.3		2.6																		
Rash	0	0	1/<1	<1/0	6.2			1.2					4	0											
Skin discolouration	0	0	<1/0																						
ALT increased	11	0	2/<1	2/0																				2.5	1.2
AST increased	6.5	0	2/0	2/0									1	4											
Total bilirubin increased	0.7	0	1/0															0	0.7						
Hypophosphataemia	0	0																							
Hypertension	3.9	0	12/0	1/0	2.1	1.1		0	3.3	0.7	9.3/0.5	0/0			0.8	0	0								
Anorexia	1.9	0	2/0	2/0		2.2		1.2	3.0	2.6	17.2/0	8.0/0	3	4						39.2	8.2				
Hyperglycaemia	0	0											11	2											

[1] Reported as % (Grade 3/Grade 4)
[2] Grade 3 or worse AEs reported.
[3] Reported during induction treatment
[4] AE of moderate to severe intensity reported at week 12
[5] Grade 4 AEs reported

Study	VEG105192		Motzer 2009[1]		Escudier 2009[2]		Target Study		AVOREN trial[2]		CALGB 90206[1]		Global ARCC trial		Negrier 2007			CRECY Trial[3]		MRC RE01[4]		Kriegmair 1995		Pyrhonen 1999[5]	
	PAZO	PLAC	SUN	IFN	SORA	IFN	SORA	PLAC	BEV + IFN	IFN	BEV + IFN	IFN	TEM	IFN	IFN	IL-2	BSC	IFN	IL-2	IFN	BSC	IFN + BSC	BSC	IFN + BSC	BSC
Hypophosphataemia			6/<1	6/0																					
Arthralgia	0	0	<1/0	<1/0																					
Flank pain	0	0																							
Altered taste	0	0	<1/0																						
Headache	0	0	1/0						2.1	1.3			1	0											
Epistaxis	0	0	1/0	13/<1																					
Infection	1.9	0											5	4				0.7	8						
Flu-like symptoms	0	0				2.2			3.0	2															
Anaemia	1.9	1.3	6/2	5/1				0	2.7	5.6	3.3/0.5	3.4/0.3	20	22	6.6	4.8	0	6.1	17.4			0	0		
Leucopenia	0	0	8/0	2/0									1	5				0.7	0.7			0	0		
Lymphocytopaenia	0	0	16/2	24/2											4.1	2.4	2.5								
Neutropaenia	1.3	0	16/2	8/1			0	0 (0)	4.5	2.3	7.9/1.1	8.3/0.3	3	7	4.1	0	0							15.2	
Thrombocytopaenia	1.9	0	8/1	1/0			0	0 (0)	2.1	1.0	1.9/0.3	0.6/0	1	0	0	0.8	0	0	3.6			0	0		
Depression	0	0			0				3.0	1.3															
Congestive heart failure	0.7	0							0.3	0					0	0	0								
Hypothyroidism	0	0	2/0	<1/0																					

Evaluable N was used to calculate %. No AE data for specific grade ¾ AE's for Steineck 1990

Dark green represents 0% patients reported AE, **Light green** represents 1-25% patients with AE, **Yellow** represents 26-50% patients with AE, **Orange** represents 51-75% patients with AE, **Red** represents 76-100% patients with AE.

PAZO = pazopanib. SUN = Sunitinib. SORA = sorafenib. BEV = bevacizumab. TEM = temsirolimus. IFN = Interferon. Bev + IFN = Bevacizumab plus interferon. BSC = Best supportive care.

5.4 Tolerability

Reasons for treatment discontinuations were reported in seven studies and active interventions demonstrated similar tolerability to comparators in all, except one study (AVOREN trial, (Escudier 2007c)). Treatment discontinuation due to death ranged from 2 patients (Escudier 2009c) to 23 patients (Motzer 2009) in the active intervention group and from 2 patients (Pyrhonen 1999) to 20 patients (Motzer 2009) in the comparator groups. In the VEG1015192 study, treatment discontinuation (all and due to death) was numerically lower in the pazopanib group compared to placebo group.

Treatment discontinuation due to AEs ranged from four patients (Pyrhonen 1999) to 95 patients (AVOREN trial, (Escudier 2007c)) in the active interventions and from 0 (Pyrhonen 1999) to 86 patients (Motzer 2009) in comparator groups.

Table 38: Summary of tolerability – treatment discontinuation

Study	Intervention	N	Treatment discontinuation % (n)			Comments
			All	Due to death	Due to AE	
VEG105192	Pazopanib	155	75.48 (117)	3.87 (6)	10.97 (17)	Treatment discontinuations are reported.
	Placebo	78	87.18 (68)	7.69 (6)	5.13 (4)	
Motzer 2009	Sunitinib	375	86.13 (323)	6.13 (23)	18.67 (70)	Number of deaths includes two patients who had crossed over from IFN-alpha group. Death reported for patients while on study; defined as death occurring on treatment up to 28 days after last dose (one death was considered treatment related).
	IFN	375	98.4 (369)	5.33 (20)	22.93 (86)	Death reported for patients while on study; defined as death occurring on treatment up to 28 days after last dose (two deaths were considered treatment related).
Escudier 2009	Sorafenib	97	15.46 (15)	4.12 (4)	11.34 (11)	Treatment discontinuations are reported.
	IFN	92	17.39 (16)	2.17 (2)	15.22 (14)	
AVOREN trial	Bevacizumab + IFN	327	32.72 (107)	2.45 (8)	29.05 (95)	All withdrawals reported for patients excluding those died or progressed during treatment (n = 151). Withdrawal due to death reported for death due to AE (not due to disease progression). Withdrawal due to AE reported for AE leading to study discontinuation due to any study drug.
	IFN	322	17.39 (56)	2.17 (7)	11.49 (37)	All withdrawals reported for patients excluding those died or progressed during treatment (n = 224). Withdrawal due to death reported for death due to AE (not due to disease progression). Withdrawal due to AE reported for AE leading to study discontinuation due to any study drug.
CALGB 90206	Bevacizumab + IFN	369	96.21 (355)		23.04 (85)	Data was reported for treatment discontinuations. Four patients withdrew from the study due to lost to follow-up (lost: one; consent withdrawal: 3).
	IFN	363	97.8 (355)		18.18 (66)	Data was reported for treatment discontinuations. Two patients withdrew from the study due to lost to follow-up.
Global ARCC trial	Temsirolimus	209	95.22 (199)	2.87 (6)	7.18 (15)	Data was reported for treatment discontinuations. Other reasons were: disease progression (n = 153); symptomatic deterioration (n = 14); patient request (n = 8); other (n = 2); and protocol violation (n = 1). HR for death in interferon alone vs. temsirolimus alone group was 0.73 (95% CI; 0.58 to 0.92); p=0.008
	IFN	207	93.72 (194)	4.83 (10)	14.01 (29)	Data was reported for treatment discontinuations. Other reasons were: disease progression (n = 115); symptomatic deterioration (n = 28); patient request (n = 6); other (n = 4); and protocol violation (n = 2). HR for death in interferon alone vs. temsirolimus alone group was 0.73 (95% CI; 0.58 to 0.92); p=0.008
Pyrhonen 1999	BSC	81	95.06 (77)	7.41 (6)	0 (0)	Withdrawals after developing PD = 69; Withdrawals due to other reasons: 2; treatment discontinuation due to death after progressive disease: 6 patients.
	IFN + BSC	79	77.22 (61)	2.53 (2)	6.33 (5)	Withdrawals after developing PD = 50; Withdrawals due to other reasons = 4; 5 patients withdrew due to AE before completing 12 months treatment and one withdrew due to AE after 12 months of treatment; treatment discontinuation due to death after progressive disease: 2 patients.

BSC = best supportive care, IFN = interferon alpha, ITT = intention to treat

5.5 Dose reductions and dose interruptions

Dose reductions were reported in nine studies, though the definition of 'dose reduction' varied between them (see Table 39). Dose reductions were highest amongst patients receiving combination therapy with bevacizumab plus IFN. The percentage of patients receiving dose reductions in active treatment groups ranged from 22.97% (Global ARCC trial, (Hudes 2007)) to 64.5% (CALGB trial, (Rini 2008a)) compared with 3.85% (VEG105192) to 48% (CALGB trial, (Rini 2008a)) in comparator groups. Dose reductions were reported in 24% of patients receiving pazopanib in VEG105192 and 50% of patients receiving sunitinib in the sunitinib trial (Motzer 2009).

Dose interruptions were reported in six studies. The definition of dose interruption varied across the studies (see Table 39). Dose interruption was highest amongst patients receiving either temsirolimus or IFN. The percentage of patients reporting dose interruptions in active treatment groups ranged from 36.59% (Kriegmair 1995) to 66% (Global ARCC trial, (Hudes 2007)) compared with 5.13% (VEG105192) to 65% (Global ARCC trial, (Hudes 2007)) in comparator groups.

The pazopanib study reported dose reductions and dose interruptions due to adverse events. Dose interruptions because of adverse events were observed in approximately one third of patients treated with pazopanib (VEG105192) and sunitinib (Motzer 2000).

Table 39: Summary of tolerability – dose reductions and dose interruptions

Study	Intervention	Number Evaluable	Dose reductions		Dose interruptions	
			% (n)	Definition	% (n)	Definition
VEG105192	Pazopanib	155	23.23 (36)	Dose reduction due to AE was reported	36.77 (57)	Dose interruption due to AE was reported
	Placebo	78	3.85 (3)		5.13 (4)	
Motzer 2009	Sunitinib	375	50 (188)		38 (143)	Dose interruption due to AE was reported
	IFN	375	27 (101)		32 (120)	
Escudier 2009	Sorafenib	97	32.99 (32)		60.82 (59)	
	IFN	92	26.09 (24)		43.48 (40)	
AVOREN trial	Bev + IFN	327	40.06 (131)	Reduction in dose of IFN from 9 MIU to 6 or 3 MIU.		
	IFN	322	30.12 (97)			
CALGB 90206	Bev + IFN	369	64.50 (238)	Dose reduction of IFN to 6 MU and to 3 MU was undertaken in 170 and 68 patients respectively.	61.79 (228)	Treatment delays of 4 to 6 days owing to toxicity occurred in 31 patients, of 7 to 9 days occurred in 51 patients and of more than 9 days occurred in 146 patients.
	IFN	363	47.66 (173)	Dose reduction of IFN to 6 MU and to 3 MU was undertaken in 136 and 37 patients respectively.	31.68 (115)	Treatment delays of 4 to 6 days owing to toxicity occurred in 24 patients, of 7 to 9 days occurred in 31 patients and of more than 9 days occurred in 60 patients.
Global ARCC trial	Temsirolimus	209	22.97 (48)	Patients with ≥1 dose reduction	65.55 (137)	Patients with ≥1 dose delay
	IFN	207	37.68 (78)		65.22 (135)	
MRC RE01#	IFN	167	24 (40)			
	BSC	168	7 (12)			
Pyrhonen 1999	IFN + BSC	79	53.16 (42)	Patients were considered to have received a reduced dose if they received doses that were less than 18 million units, including missed treatments, for more than 6 consecutive days.		Treatment was prolonged beyond 12 months for six patients and was later stopped due to AEs (one patient), death (one patient), or development of PD (four patients).
	BSC					Treatment was prolonged beyond 12 months for one patient and was later stopped after development of PD.
Kriegmair 1995	BSC + IFN	41	26.83 (11)	Dose was reduced to 3 to 6 million units in these patients.	36.59 (15)	Dose was interrupted at least intermittently in these patients. Mean duration of interruption was 4.3 weeks (range: 1 to 14 weeks).
	BSC					

AE = adverse event BEV = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha, ITT = intention to treat; #extracted from Ritchie 1999 as updated results were not reported in Hancock 2000

6 Meta-analysis of results from studies in the clinical systematic review

This section presents the results of the meta-analysis. Additionally, the results are presented where this methodology was used to calculate the effect size and confidence interval along with the p values for a single study (See section 3.1.6). Statistical heterogeneity is presented by means of I^2 statistics.

For tables in this section, the highlighted cells indicate statistically significant results (blue favouring intervention and yellow favouring control group).

6.1 Efficacy outcomes

6.1.1 Progression free survival

PFS at one year was reported in two studies, only one of which showed a statistically significant difference, favouring IFN compared to sorafenib (Table 40).

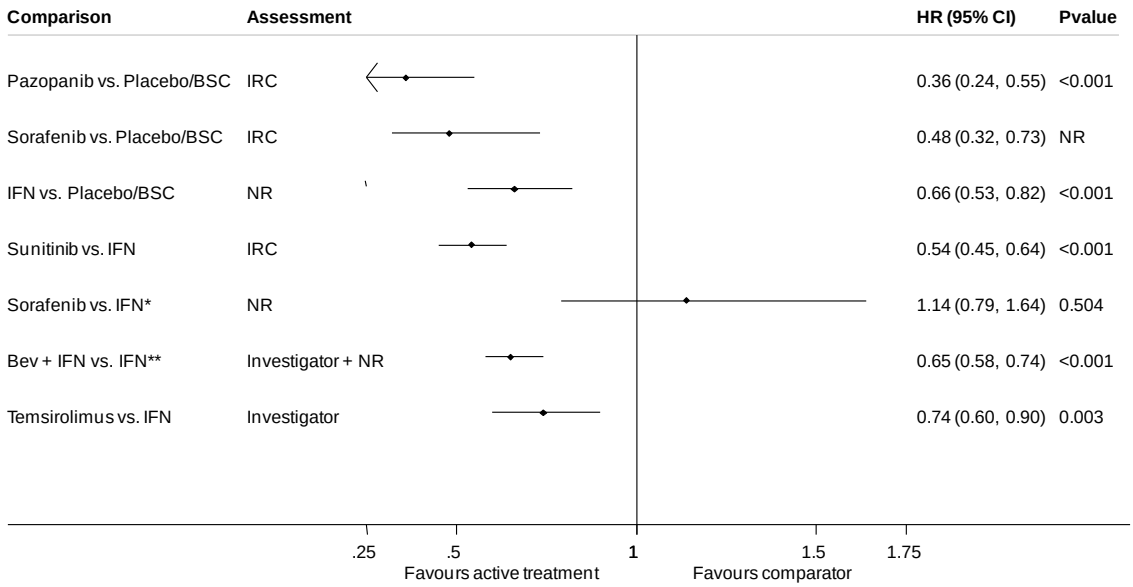
Table 40: Results of meta-analysis – PFS at 1 year

Comparison	Studies	N	Fixed effects		Random effects		I2
			RR (95% CI)	P value	RR (95% CI)	P value	
Sorafenib versus IFN	1	189	0.37 (0.2, 0.7)	<0.001	0.37 (0.2, 0.7)	<0.001	0
IL-2 versus IFN	1	285	1.24 (0.69, 2.23)	0.47	1.24 (0.69, 2.23)	0.47	0

RR = relative risk, IFN = interferon alpha, CI = confidence interval.

Eight of the included studies reported HRs for PFS (Figure 12) for one of the interventions of interest versus one of the comparators of interest. All active interventions showed improved efficacy (favourable HR) over placebo/BSC and the difference was statistically significant. Targeted therapies (sunitinib, bevacizumab plus IFN, temsirolimus, sorafenib) showed favourable hazard ratios compared to IFN, with the exception of sorafenib, where the difference was in favour of IFN (although this difference was not statistically significant).

Figure 12: Forest plot presenting the hazard ratios for PFS



*Inversed HR reported, compared to that in the publication. **Calculated by pooling results from 2 studies in which the author reported hazard ratios (fixed effect). Inv = investigator, NR = not reported, IRC= independent review committee, BSC = best supportive care, IFN = interferon alpha, HR = hazard ratio.

6.1.2 Time to progression

TTP was reported in two studies, out of which one showed statistically significant difference, favouring bevacizumab plus IFN compared to IFN alone.

Table 41: HR for TTP as reported by authors in the included studies

Study	Intervention	Comparator	HR (95% CI)	P value
Escudier 2009*	Sorafenib	IFN	1.12 (0.78, 1.64)	>0.05
AVOREN trial	Bev + IFN	IFN	0.61 (0.51, 0.73)	0.0001

*Inversed HR reported, compared to that in the publication.
Bev = bevacizumab, IFN = Interferon Alpha, HR = Hazard ratio, CI = Confidence Interval.

6.1.3 OS

In addition to pazopanib study (VEG105192), only three studies (MRC RE01, Pyrhonen 1999, Kriegmair 1995) comparing IFN with placebo/BSC reported sufficient data for meta-analysis of the rate of OS. Overall survival rate at 1 year was numerically higher in the pazopanib group compared to placebo however the difference did not reach statistical significance.

The OS rate was statistically significant and superior for IFN compared to placebo/BSC at all time points assessed (Table 42). Overall survival at study endpoint (5-year) was reported in only one study (Pyrhonen 1999).

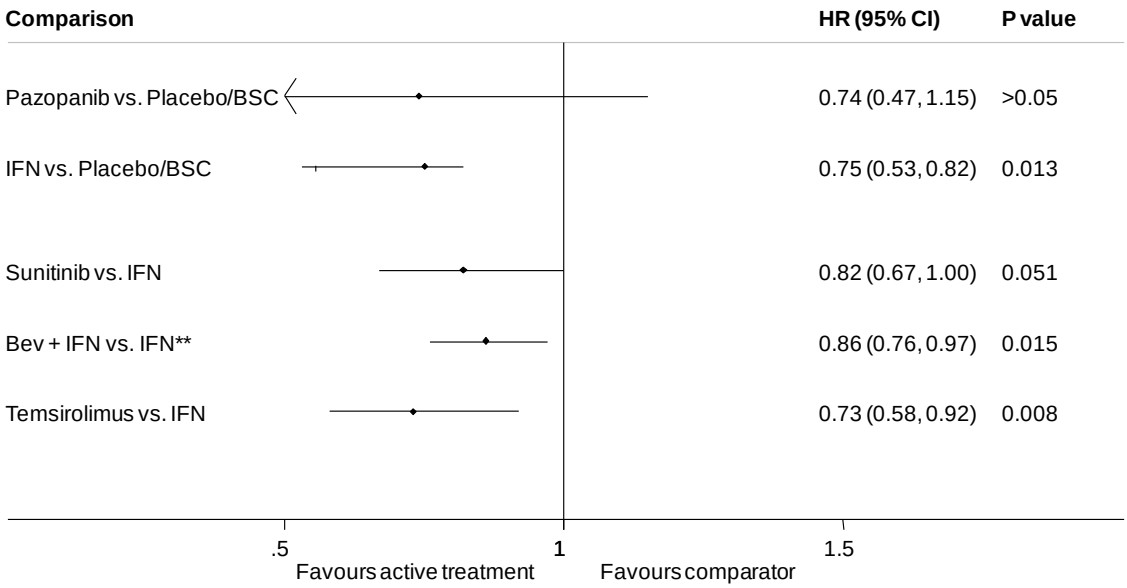
Table 42: Result of meta-analysis – OS rates

Outcome	Intervention	Comparator	Studies	N	Fixed effects		Random effects		I2
					RR (95% CI)	P value	RR (95% CI)	P value	
OS at 1 year	Pazopanib	Placebo/BSC	1	233	1.21 (0.98, 1.5)	0.07	1.21 (0.98, 1.5)	0.07	0
OS at 1 year	IFN	Placebo/BSC	2	510	1.36 (1.09, 1.68)	0.01	1.36 (1.1, 1.69)	0.01	0
OS at 2 year	IFN	Placebo/BSC	2	439	1.55 (1.02, 2.37)	0.04	1.55 (1.02, 2.37)	0.04	0
OS at endpoint	IFN	Placebo/BSC	1	160	7.18 (0.38, 136.69)	0.190	7.18 (0.38, 136.69)	0.190	0

RR = Relative Risk, CI = Confidence Interval

Six of the included studies provided HRs for OS (Figure 13) for one of the interventions of interest versus one of the comparators of interest. The hazard ratio for OS demonstrated an increased survival advantage with pazopanib compared to placebo; however, this difference did not reach statistical significance. Targeted therapy (sunitinib, bevacizumab plus IFN, temsirolimus) showed favourable hazard ratio compared to IFN however.

Figure 13: Forest plot presenting the HR for OS



Bev = bevacizumab, BSC = Best Supportive Care, CI= Confidence interval, IFN = Interferon Alpha, HR = Hazard ratio.

All hazard ratios are reported by the authors of the publications. **For Bev + IFN vs. IFN comparison, HR was calculated by pooling results from 2 studies in which the author reported hazard ratios (fixed effect).

6.1.4 Response rate

The results of meta-analyses for response rates are presented in Table 43, Figure 14 and Figure 15. Overall response rate (ORR) was higher for pazopanib, sorafenib, IFN and IL-2 compared to placebo/BSC and all improvements over placebo were statistically significant, with the exception of IL-2. Sunitinib and bevacizumab plus IFN demonstrated a significantly higher ORR compared to IFN alone. A higher response rate was observed for temsirolimus over IFN; however the difference did not reach statistical significance. Sorafenib and IL-2 both showed lower ORR than IFN; however, this difference was also not statistically significant.

Pazopanib showed better efficacy than placebo for overall, complete and partial response. Sunitinib showed better efficacy than IFN for complete and partial responses. Sorafenib showed improved efficacy over placebo but not IFN therapy, although the difference did not reach statistical significance. Bevacizumab plus IFN showed superior efficacy compared to IFN alone for partial response.

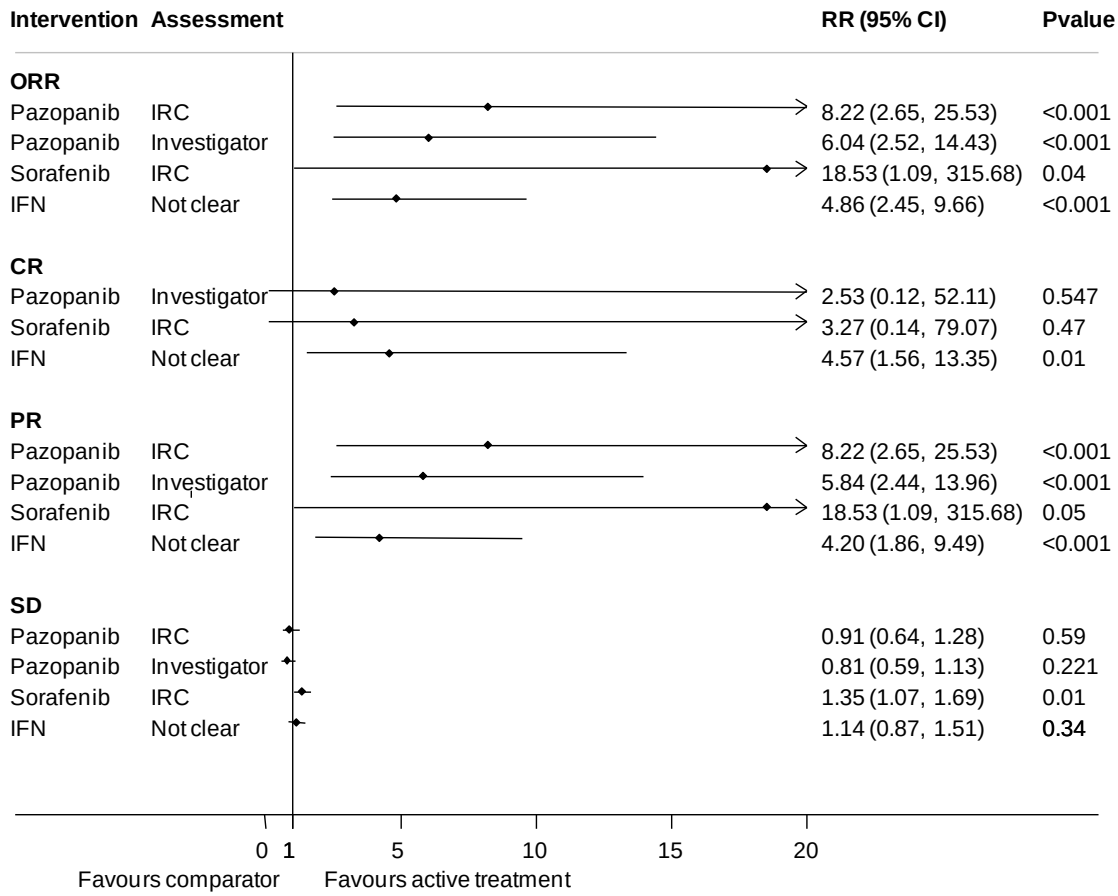
Table 43: Result of meta-analysis – response rate

Outcome	Intervention	Comparator	Studies	N	Fixed effects		Random effects		I2
					RR (95% CI)	P value	RR (95% CI)	P value	
Overall response rate	**Independent review committee**								
	Pazopanib	Placebo/BSC	1	233	8.22 (2.65, 25.53)	<0.001	8.22 (2.65, 25.53)	<0.001	0
	Sorafenib	Placebo/BSC	1	161	18.53 (1.09, 315.68)	0.044	18.53 (1.09, 315.68)	0.044	0
	Sunitinib	IFN	1	750	5.15 (3.26, 8.13)	<0.001	5.15 (3.26, 8.13)	<0.001	0
	Temsirolimus	IFN	1	416	1.78 (0.84, 3.77)	0.130	1.78 (0.84, 3.77)	0.130	0

Outcome	Intervention	Comparator	Studies	N	Fixed effects		Random effects		I2
					RR (95% CI)	P value	RR (95% CI)	P value	
	Sorafenib	IFN	1	189	0.59 (0.2, 1.75)	0.343	0.59 (0.2, 1.75)	0.343	0
	IL-2	IFN	1	285	0.87 (0.37, 2.04)	0.751	0.87 (0.37, 2.04)	0.751	0
	Investigator assessed								
	Pazopanib	Placebo/BSC	1	233	6.04 (2.52, 14.43)	<0.001	6.04 (2.52, 14.43)	<0.001	0
	Sunitinib	IFN	1	750	3.83 (2.86, 5.12)	<0.001	3.83 (2.86, 5.12)	<0.001	0
	Bev + IFN	IFN	1	649	2.55 (1.81, 3.61)	<0.001	2.55 (1.81, 3.61)	<0.001	0
	Temsirolimus	IFN	1	416	1.11 (0.58, 2.13)	0.743	1.11 (0.58, 2.13)	0.743	0
	Not clear								
	IFN	Placebo/BSC	5	904	4.86 (2.45, 9.66)	<0.001	4.27 (2.11, 8.63)	<0.001	0
Complete response	**Independent review committee**								
	Sorafenib	Placebo/BSC	1	161	3.27 (0.14, 79.07)	0.466	3.27 (0.14, 79.07)	0.466	0
	Sorafenib	IFN	1	189	0.32 (0.01, 7.67)	0.479	0.32 (0.01, 7.67)	0.479	0
	IL-2	IFN	1	285	5.32 (0.26, 109.91)	0.279	5.32 (0.26, 109.91)	0.279	0
	Investigator assessed								
	Pazopanib	Placebo/BSC	1	233	2.53 (0.12, 52.11)	0.547	2.53 (0.12, 52.11)	0.547	
	Sunitinib	IFN	1	750	2.75 (0.88, 8.56)	0.081	2.75 (0.88, 8.56)	0.081	0
	Bev + IFN	IFN	1	649	0.66 (0.19, 2.3)	0.51	0.66 (0.19, 2.3)	0.51	0
	Not clear								
	IFN	Placebo/BSC	5	904	4.57 (1.56, 13.35)	0.01	4.1 (1.33, 12.59)	0.01	0
Partial response	**Independent review committee**								
	Pazopanib	Placebo/BSC	1	233	8.22 (2.65, 25.53)	<0.001	8.22 (2.65, 25.53)	<0.001	0
	Sorafenib	Placebo/BSC	1	161	16.35 (0.95, 281.51)	0.054	16.35 (0.95, 281.51)	0.054	0
	Sunitinib	IFN	1	750	5.15 (3.26, 8.13)	<0.001	5.15 (3.26, 8.13)	<0.001	0
	Sorafenib	IFN	1	189	0.68 (0.22, 2.06)	0.492	0.68 (0.22, 2.06)	0.492	0
	IL-2	IFN	1	285	0.68 (0.27, 1.7)	0.407	0.68 (0.27, 1.7)	0.407	0
	Investigator assessed								
	Pazopanib	Placebo/BSC	1	233	5.84 (2.44, 13.96)	<0.001	5.84 (2.44, 13.96)	<0.001	0
	Sunitinib	IFN	1	750	3.93 (2.89, 5.34)	<0.001	3.93 (2.89, 5.34)	<0.001	0
	Bev + IFN	IFN	1	649	2.92 (2, 4.26)	<0.001	2.92 (2, 4.26)	<0.001	0
	Not clear								
	IFN	Placebo/BSC	5	904	4.2 (1.86, 9.49)	<0.001	3.75 (1.62, 8.66)	<0.001	0
Stable disease	**Independent review committee**								
	Pazopanib	Placebo/BSC	1	233	0.91 (0.64, 1.28)	0.587	0.91 (0.64, 1.28)	0.587	0
	Sorafenib	Placebo/BSC	1	161	1.35 (1.07, 1.69)	0.011	1.35 (1.07, 1.69)	0.011	0
	Sunitinib	IFN	1	750	1 (0.85, 1.18)	1.000	1 (0.85, 1.18)	1.000	0
	Sorafenib	IFN	1	189	1.34 (1.08, 1.66)	0.009	1.34 (1.08, 1.66)	0.009	0
	IL-2	IFN	1	285	0.69 (0.47, 1.03)	0.072	0.69 (0.47, 1.03)	0.072	0
	Investigator assessed								
	Pazopanib	Placebo/BSC	1	233	0.81 (0.59, 1.13)	0.221	0.81 (0.59, 1.13)	0.221	0
	Sunitinib	IFN	1	750	0.74 (0.64, 0.87)	<0.001	0.74 (0.64, 0.87)	<0.001	0
	Bev + IFN	IFN	1	649	0.96 (0.81, 1.15)	0.68	0.96 (0.81, 1.15)	0.68	0
	Not clear								
	IFN	Placebo/BSC	4	815	1.14 (0.87, 1.51)	0.34	1.1 (0.84, 1.44)	0.5	0

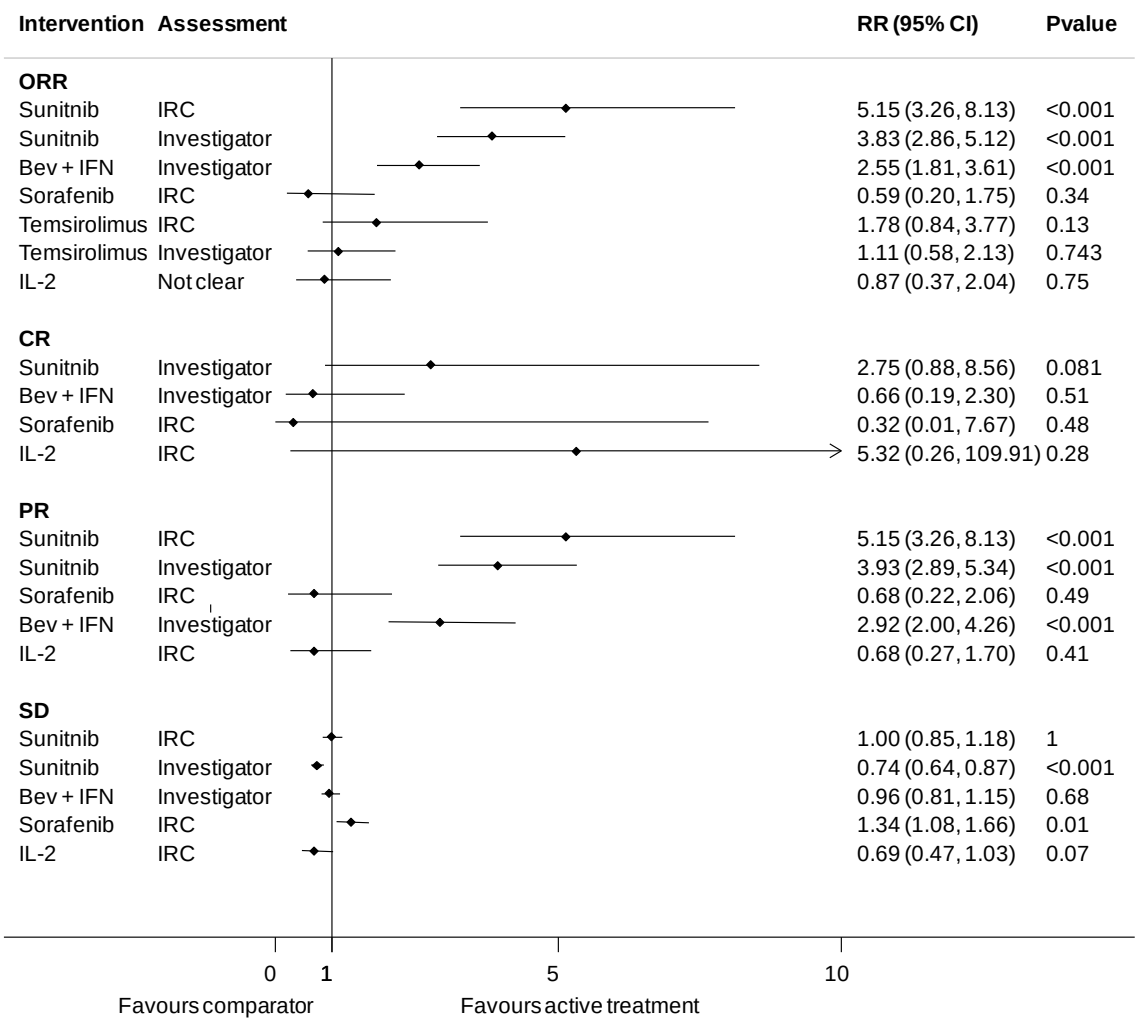
Bev = Bevacizumab, BSC = Best Supportive Care, IFN = Interferon Alpha, RR = Relative Risk, IRC = Independent review committee.

Figure 14: Forest plot for response rates in placebo/BSC controlled trials (fixed effects)



BSC = best supportive care, CR = complete response, ORR = overall response rate, IFN = interferon alpha, IL-2 = interleukin-2, PR = partial response, RR = relative risk, SD = stable disease, IRC = Independent review committee

Figure 15: Forest plot for response rates in IFN controlled trials (fixed effects)



ORR = overall response rate, CR = complete response, PR = partial response, SD = stable disease, RR = relative risk, IRC = Independent review committee

6.1.5 Time to, and duration of, response

None of the included studies provided sufficient data for meta-analysis of these outcomes. A qualitative summary is provided in section 5.1.5.

6.2 Health related quality of life

Only three of the included studies (one being the pazopanib trial) reported data on HRQoL outcomes, although none provided sufficient data for meta-analysis. A description of the reported data is presented in section 5.2.

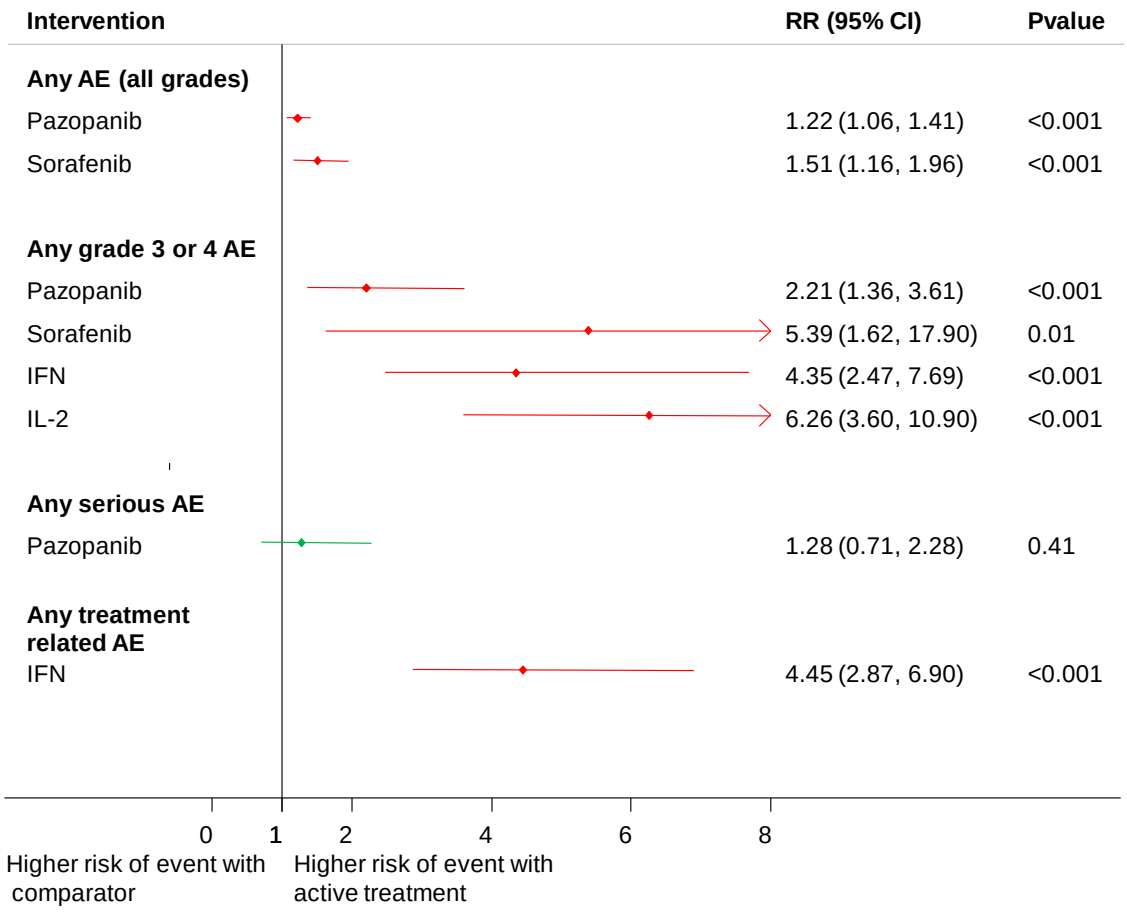
6.3 AEs

6.3.1 Any AE

Meta-analysis was performed for studies reporting the percentage of patients who experienced any AE (at all grades). Due to limited data available, pooling of results from multiple studies was not possible.

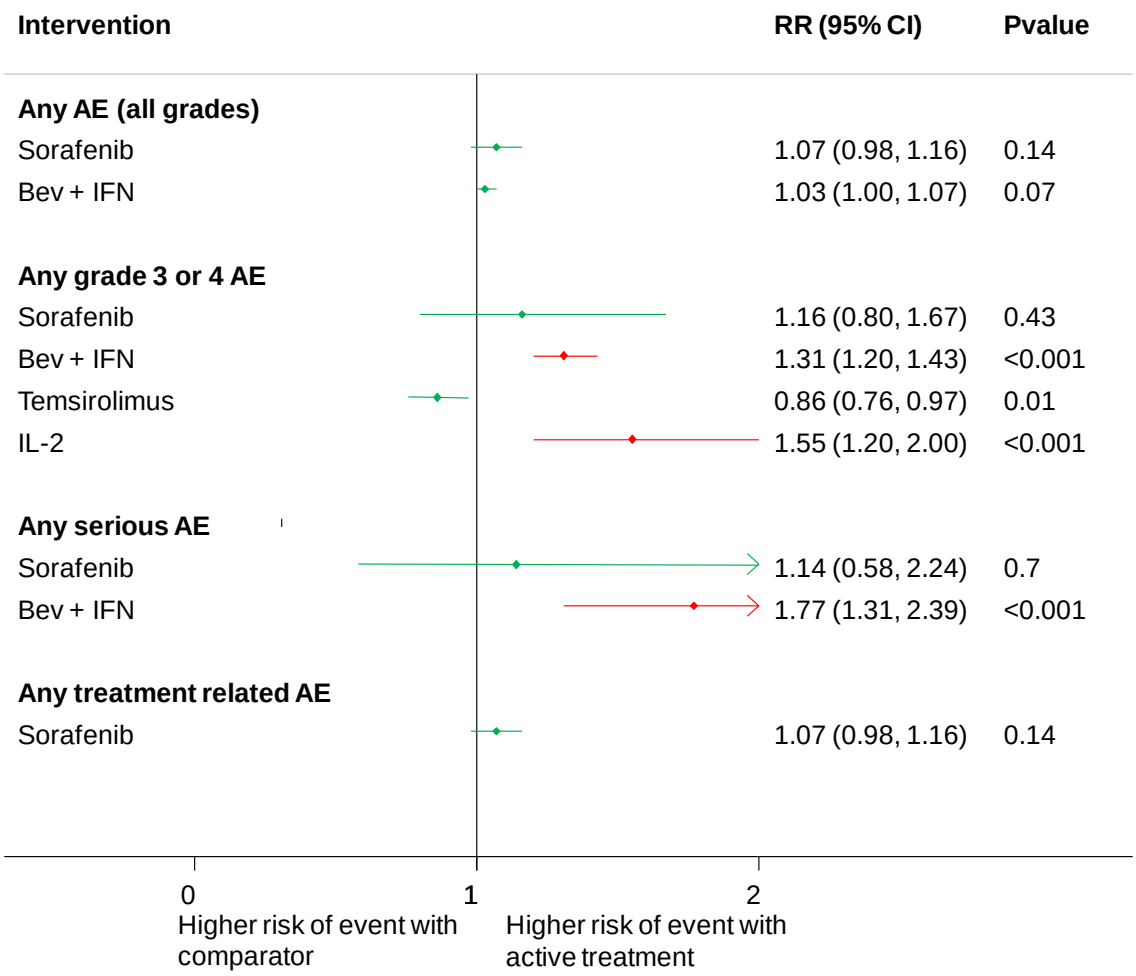
The results are presented in Figure 16 and Figure 17. Active treatments were associated with a significantly higher risk of any AE or any grade 3 or 4 AE compared to placebo/BSC. Pazopanib was, however, not associated with a significantly increased risk of any serious AE compared to placebo. Sorafenib and bevacizumab plus IFN were associated with higher risk of any serious AE compared to IFN. Temsirolimus was associated with a significantly lower risk of any grade 3 or 4 AE compared to IFN, while IL-2 and bevacizumab plus IFN were associated with significantly increased risk. On qualitative analysis, it was observed that higher proportion of patients had increased AST and ALT levels after the treatment with pazopanib as compared with sunitinib.

Figure 16: Forest plot for any AE in placebo/BSC controlled trials (fixed effects)



AE = Adverse event, IFN = interferon alpha, IL-2 = interleukin-2. Those highlighted in red are statistically significant ($p<0.05$), those in green are not statistically significant ($p>0.05$).

Figure 17: Forest plot for any AE in IFN controlled trials (fixed effects)



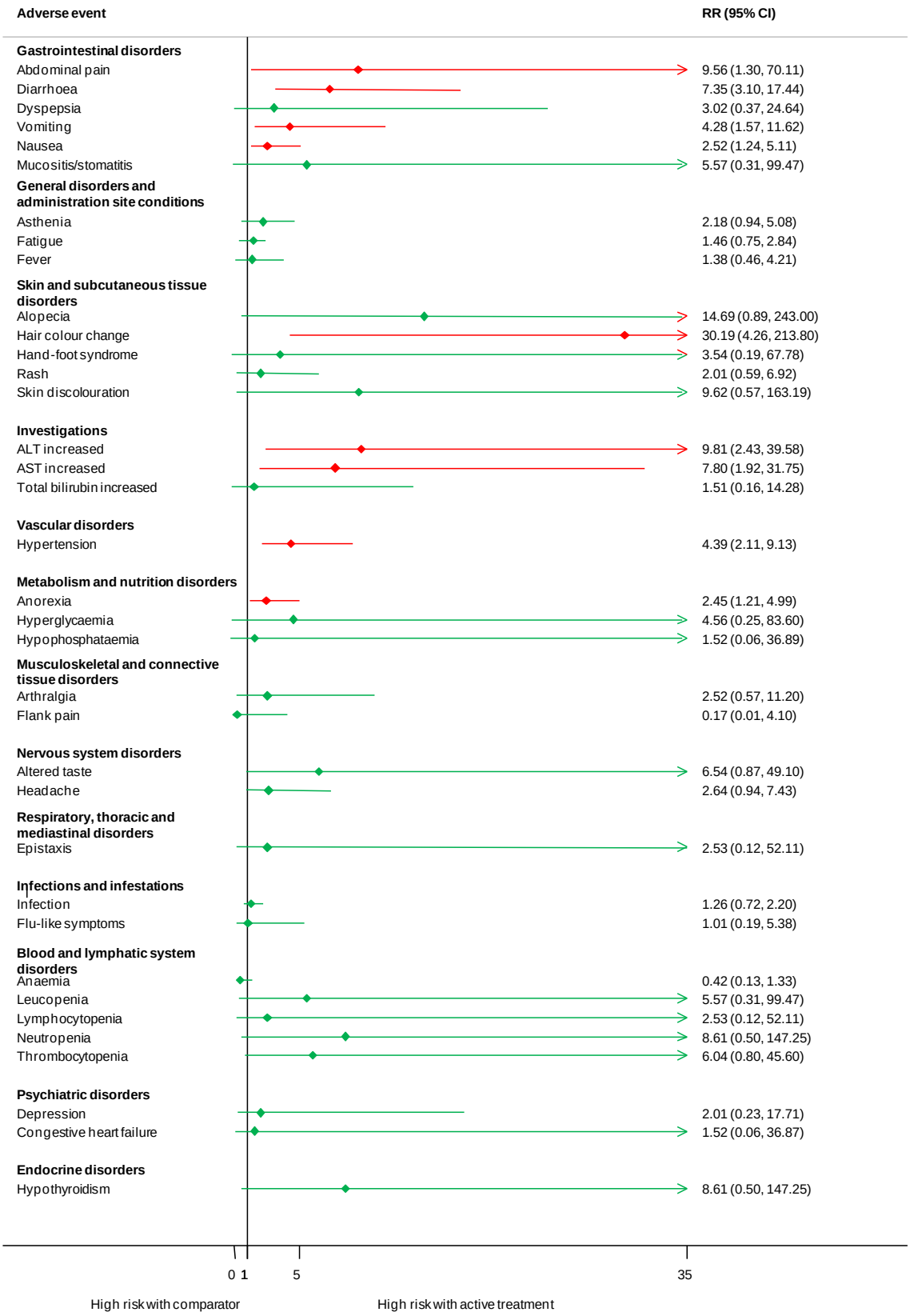
AE = Adverse Event, BEV = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha, ITT = intention to treat. Those highlighted in red are statistically significant ($p<0.05$), those in green are not statistically significant ($p>0.05$) or favours comparator ($p<0.05$).

6.3.2 Specific AEs

Meta-analysis was performed for studies reporting the percentage of patients experiencing specific AEs at all grades. The results are presented in Figure 18, Table 44 and Table 45.

Although, the effect size demonstrated a numerically higher risk of specific AEs with pazopanib than placebo/BSC, the difference rarely reached statistical significance, suggesting that pazopanib is a well tolerated treatment, Figure 18. The only AEs for which pazopanib demonstrated an increased risk over placebo/BSC were abdominal pain, hypertension, diarrhoea, vomiting and nausea, hair colour change, increased ALT, increased AST and anorexia.

Figure 18: Specific AE results from meta-analysis of pazopanib versus placebo/BSC



Those highlighted in red are statistically significant ($p<0.05$), those in green are not statistically significant ($p>0.05$).

Sorafenib and IFN were also compared to placebo/BSC for the risk of specific AEs, Table 44. Sorafenib was associated with significantly increased risk of diarrhoea, alopecia,

hand-foot syndrome and rash compared to placebo/BSC, while IFN was associated with significantly increased risk of fatigue, flu-like symptoms and leucopenia compared to placebo/BSC.

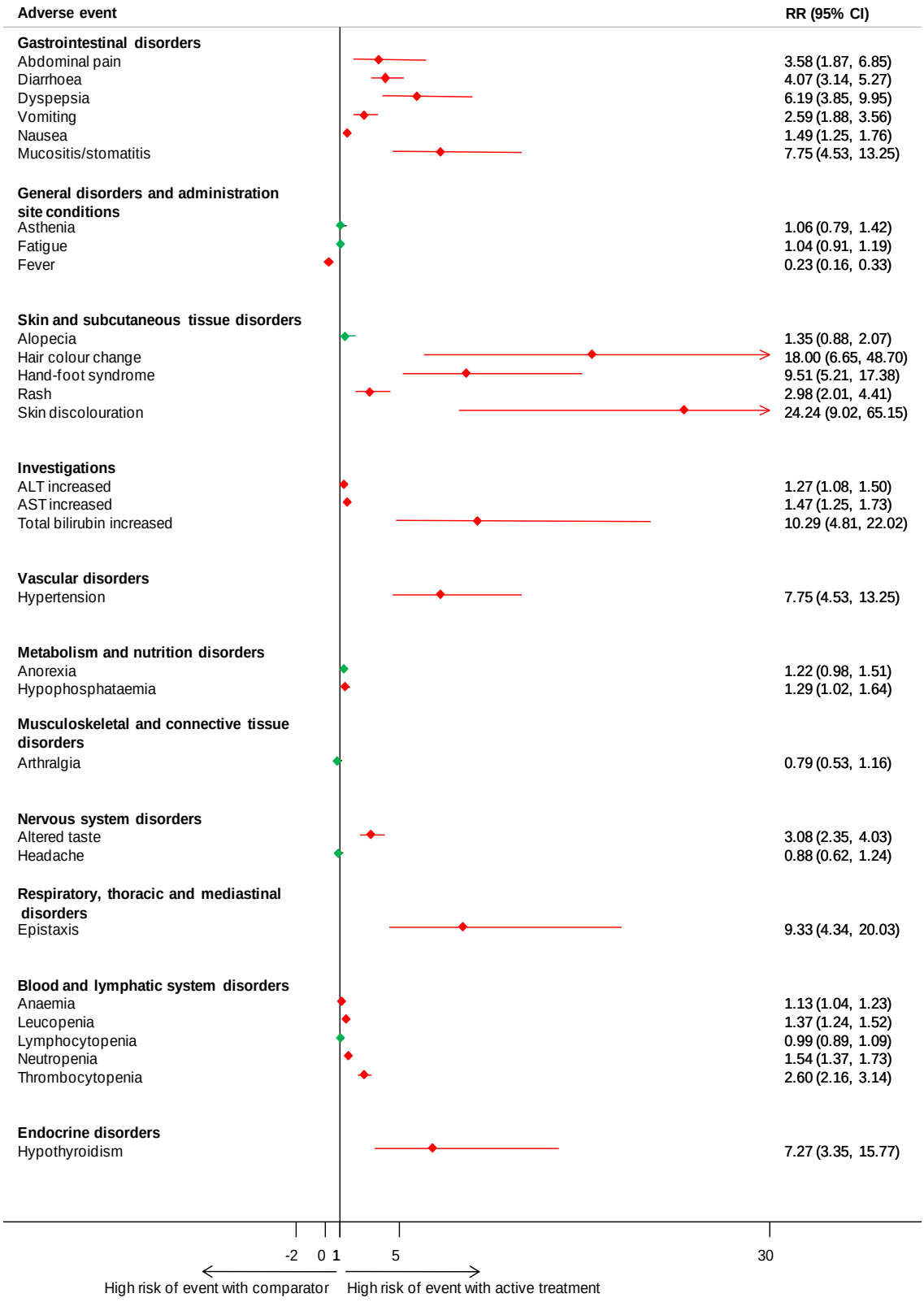
Table 44: Result of meta-analysis – AEs (all grades) versus placebo

Outcome	Intervention	Studies	N	Fixed effects		Random effects		I2
				RR (95% CI)	P value	RR (95% CI)	P value	
Gastrointestinal disorders								
Abdominal pain	Pazopanib	1	233	9.56 (1.3, 70.11)	0.03	9.56 (1.3, 70.11)	0.03	0
Diarrhoea	Pazopanib	1	233	7.35 (3.1, 17.44)	<0.001	7.35 (3.1, 17.44)	<0.001	0
	Sorafenib	1	160	2.52 (1.23, 5.15)	0.01	2.52 (1.23, 5.15)	0.01	0
	IFN	1	60	3 (0.13, 70.83)	0.5	3 (0.13, 70.83)	0.5	0
Dyspepsia	Pazopanib	1	233	3.02 (0.37, 24.64)	0.3	3.02 (0.37, 24.64)	0.3	0
Vomiting	Pazopanib	1	233	4.28 (1.57, 11.62)	<0.001	4.28 (1.57, 11.62)	<0.001	0
	Sorafenib	1	160	4.31 (0.49, 37.73)	0.19	4.31 (0.49, 37.73)	0.19	0
Nausea	Pazopanib	1	233	2.52 (1.24, 5.11)	0.01	2.52 (1.24, 5.11)	0.01	0
	Sorafenib	1	160	0.9 (0.41, 1.96)	0.79	0.9 (0.41, 1.96)	0.79	0
	IFN	1	60	3.00 (0.13, 70.83)	0.49	3.00 (0.13, 70.83)	0.49	0
Mucositis/stomatitis	Pazopanib	1	233	5.57 (0.31, 99.47)	0.24	5.57 (0.31, 99.47)	0.24	0
	Sorafenib	1	160	9.69 (0.53, 177.11)	0.13	9.69 (0.53, 177.11)	0.13	0
General disorders and administration site conditions								
Asthenia	Pazopanib	1	233	2.18 (0.94, 5.08)	0.07	2.18 (0.94, 5.08)	0.07	0
Fatigue	Pazopanib	1	233	1.46 (0.75, 2.84)	0.27	1.46 (0.75, 2.84)	0.27	0
	Sorafenib	1	160	1.86 (0.95, 3.66)	0.07	1.86 (0.95, 3.66)	0.07	0
	IFN	1	60	6.68 (2.21, 20.09)	0.001	6.68 (2.21, 20.09)	0.001	0
Fever	Pazopanib	1	233	1.38 (0.46, 4.21)	0.57	1.38 (0.46, 4.21)	0.57	0
Skin and subcutaneous tissue disorders								
Alopecia	Pazopanib	1	233	14.69 (0.89, 243)	0.06	14.69 (0.89, 243)	0.06	0
	Sorafenib	1	160	8.26 (2.58, 26.43)	<0.001	8.26 (2.58, 26.43)	<0.001	0
	IFN	1	60	3 (0.13, 70.83)	0.5	3 (0.13, 70.83)	0.5	0
Hair colour change	Pazopanib	1	233	30.19 (4.26, 213.8)	<0.001	30.19 (4.26, 213.8)	<0.001	0
Hand-foot syndrome	Pazopanib	1	233	3.54 (0.19, 67.78)	0.4	3.54 (0.19, 67.78)	0.4	0
	Sorafenib	1	160	6.11 (1.86, 20.03)	<0.001	6.11 (1.86, 20.03)	<0.001	0
Rash	Pazopanib	1	233	2.01 (0.59, 6.92)	0.27	2.01 (0.59, 6.92)	0.27	0
	Sorafenib	1	160	2.59 (1.32, 5.05)	0.01	2.59 (1.32, 5.05)	0.01	0
	IFN	1	60	3 (0.33, 27.23)	0.33	3 (0.33, 27.23)	0.33	0
Skin discolouration	Pazopanib	1	233	9.62 (0.57, 163.19)	0.12	9.62 (0.57, 163.19)	0.12	0
Investigations								
ALT increased	Pazopanib	1	233	9.81 (2.43, 39.58)	<0.001	9.81 (2.43, 39.58)	<0.001	0
AST increased	Pazopanib	1	233	7.8 (1.92, 31.75)	<0.001	7.8 (1.92, 31.75)	<0.001	0
Total bilirubin increased	Pazopanib	1	233	1.51 (0.16, 14.28)	0.72	1.51 (0.16, 14.28)	0.72	0
	IFN	1	60	3 (0.13, 70.83)	0.5	3 (0.13, 70.83)	0.5	0
Vascular disorders								
Hypertension	Pazopanib	1	233	4.39 (2.11, 9.13)	<0.001	4.39 (2.11, 9.13)	<0.001	0
	Sorafenib	1	160	11.85 (0.67, 210.74)	0.09	11.85 (0.67, 210.74)	0.09	0
	IFN	1	60	0.21 (0.01, 4.26)	0.31	0.21 (0.01, 4.26)	0.31	0
Metabolism and nutrition disorders								
Anorexia	Pazopanib	1	233	2.45 (1.21, 4.99)	0.01	2.45 (1.21, 4.99)	0.01	0
	Sorafenib	1	160	1.23 (0.47, 3.24)	0.67	1.23 (0.47, 3.24)	0.67	0
	IFN	1	60	6.00 (0.77, 46.87)	0.08	6.00 (0.77, 46.87)	0.08	0
Hyperglycaemia	Pazopanib	1	233	4.56 (0.25, 83.6)	0.31	4.56 (0.25, 83.6)	0.31	0
Hypophosphataemia	Pazopanib	1	233	1.52 (0.06, 36.89)	0.797	1.52 (0.06, 36.89)	0.797	0
Musculoskeletal and connective tissue disorders								
Arthralgia	Pazopanib	1	233	2.52 (0.57, 11.2)	0.23	2.52 (0.57, 11.2)	0.23	0
Flank pain	Pazopanib	1	233	0.17 (0.01, 4.1)	0.27	0.17 (0.01, 4.1)	0.27	0
Nervous system disorders								
Altered taste	Pazopanib	1	233	6.54 (0.87, 49.1)	0.07	6.54 (0.87, 49.1)	0.07	0
Headache	Pazopanib	1	233	2.64 (0.94, 7.43)	0.07	2.64 (0.94, 7.43)	0.07	0
Respiratory, thoracic and mediastinal disorders								
Epistaxis	Pazopanib	1	233	2.53 (0.12, 52.11)	0.55	2.53 (0.12, 52.11)	0.55	0
	IFN	1	60	3 (0.13, 70.83)	0.5	3 (0.13, 70.83)	0.5	0

Outcome	Intervention	Studies	N	Fixed effects		Random effects		I2
				RR (95% CI)	P value	RR (95% CI)	P value	
Infections and infestations								
Infection	Pazopanib	1	233	1.26 (0.72, 2.2)	0.42	1.26 (0.72, 2.2)	0.42	0
Flu-like symptoms	Pazopanib	1	233	1.01 (0.19, 5.38)	0.99	1.01 (0.19, 5.38)	0.99	0
	IFN	1	60	20.33 (4.27, 96.93)	<0.001	20.33 (4.27, 96.93)	<0.001	0
Blood and lymphatic system disorders								
Anaemia	Pazopanib	1	233	0.42 (0.13, 1.33)	0.14	0.42 (0.13, 1.33)	0.14	0
	Sorafenib	1	160	3.23 (0.13, 78.14)	0.47	3.23 (0.13, 78.14)	0.47	0
	IFN	1	60	1.33 (0.66, 2.69)	0.42	1.33 (0.66, 2.69)	0.42	0
Leucopenia	Pazopanib	1	233	5.57 (0.31, 99.47)	0.24	5.57 (0.31, 99.47)	0.24	0
	IFN	1	60	29 (1.81, 465.07)	0.02	29 (1.81, 465.07)	0.02	0
Lymphocytopenia	Pazopanib	1	233	2.53 (0.12, 52.11)	0.55	2.53 (0.12, 52.11)	0.55	0
Neutropenia	Pazopanib	1	233	8.61 (0.5, 147.25)	0.14	8.61 (0.5, 147.25)	0.14	0
Thrombocytopenia	Pazopanib	1	233	6.04 (0.8, 45.6)	0.08	6.04 (0.8, 45.6)	0.08	0
	IFN	1	60	5 (0.25, 99.95)	0.29	5 (0.25, 99.95)	0.29	0
Psychiatric disorders								
Depression	Pazopanib	1	233	2.01 (0.23, 17.71)	0.53	2.01 (0.23, 17.71)	0.53	0
Congestive heart failure	Pazopanib	1	233	1.52 (0.06, 36.87)	0.8	1.52 (0.06, 36.87)	0.8	0
Endocrine disorders								
Hypothyroidism	Pazopanib	1	233	8.61 (0.50, 147.25)	0.137	8.61 (0.50, 147.25)	0.137	0

For many of the specific AEs, sunitinib was associated with a greater risk of patients experiencing the event than IFN, the majority of which were statistically significant, Figure 19. AEs within the GI disorders, blood and lymphatic disorders, skin and subcutaneous tissue disorders, investigations classes and endocrine disorders, generally all demonstrated a significantly increased risk with sunitinib. One exception was observed in this trend; IFN was associated with a significantly increased risk of fever compared to sunitinib.

Figure 19: Specific AE results from meta-analysis of sunitinib versus IFN



Those highlighted in red are statistically significant ($p<0.05$), those in green are not statistically significant ($p>0.05$).

Sorafenib was also associated with an increased risk of certain AEs compared to IFN, in particular, diarrhoea, alopecia, hand-foot syndrome and rash, as was observed with the sorafenib versus placebo comparison, Table 45. Sorafenib was, however, associated with

a significantly lower risk of fever, hypertension and depression than IFN. Bevacizumab in combination with IFN was associated with an increased risk of several specific AEs including but not limited to thrombocytopenia, compared to IFN; however, none of the differences reached statistical significance. Lastly, temsirolimus was associated with an increased risk of vomiting, mucositis/stomatitis, rash and infection compared to IFN.

Table 45: Result of meta-analysis – AEs (all grades) versus IFN

Outcome	Intervention	Studies	N	Fixed effects		Random effects		I2
				RR (95% CI)	P value	RR (95% CI)	P value	
Gastrointestinal disorders								
Abdominal pain	Sunitinib	1	735	3.58 (1.87, 6.85)	<0.001	3.58 (1.87, 6.85)	<0.001	0
	Sorafenib	1	187	1.86 (0.58, 5.95)	0.3	1.86 (0.58, 5.95)	0.3	0
	Temsirolimus	1	408	1.24 (0.83, 1.86)	0.29	1.24 (0.83, 1.86)	0.29	0
Diarrhoea	Sunitinib	1	735	4.07 (3.14, 5.27)	<0.001	4.07 (3.14, 5.27)	<0.001	0
	Sorafenib	1	187	4.47 (2.5, 8.01)	<0.001	4.47 (2.5, 8.01)	<0.001	0
	Bev + IFN	1	641	1.32 (0.95, 1.85)	0.1	1.32 (0.95, 1.85)	0.1	0
	Temsirolimus	1	408	1.35 (0.94, 1.92)	0.1	1.35 (0.94, 1.92)	0.1	0
Dyspepsia	Sunitinib	1	735	6.19 (3.85, 9.95)	<0.001	6.19 (3.85, 9.95)	<0.001	0
Vomiting	Sunitinib	1	735	2.59 (1.88, 3.56)	<0.001	2.59 (1.88, 3.56)	<0.001	0
	Sorafenib	1	187	0.93 (0.45, 1.89)	0.84	0.93 (0.45, 1.89)	0.84	0
	Temsirolimus	1	408	0.69 (0.48, 0.98)	0.04	0.69 (0.48, 0.98)	0.04	0
Nausea	Sunitinib	1	735	1.49 (1.25, 1.76)	<0.001	1.49 (1.25, 1.76)	<0.001	0
	Sorafenib	1	187	0.67 (0.39, 1.14)	0.14	0.67 (0.39, 1.14)	0.14	0
	Temsirolimus	1	408	0.9 (0.71, 1.15)	0.41	0.9 (0.71, 1.15)	0.41	0
Mucositis/stomatitis	Sunitinib	1	735	7.75 (4.53, 13.25)	<0.001	7.75 (4.53, 13.25)	<0.001	0
	Temsirolimus	1	408	5.05 (2.43, 10.48)	<0.001	5.05 (2.43, 10.48)	<0.001	0
General disorders and administration site conditions								
Asthenia	Sunitinib	1	735	1.06 (0.79, 1.42)	0.7	1.06 (0.79, 1.42)	0.7	0
	Bev + IFN	1	641	1.17 (0.92, 1.49)	0.2	1.17 (0.92, 1.49)	0.2	0
	Temsirolimus	1	408	0.8 (0.67, 0.94)	0.01	0.8 (0.67, 0.94)	0.01	0
Fatigue	Sunitinib	1	735	1.04 (0.91, 1.19)	0.55	1.04 (0.91, 1.19)	0.55	0
	Sorafenib	1	187	1 (0.72, 1.39)	1	1 (0.72, 1.39)	1	0
	Bev + IFN	1	641	1.2 (0.94, 1.52)	0.14	1.2 (0.94, 1.52)	0.14	0
Fever	Sunitinib	1	735	0.23 (0.16, 0.33)	<0.001	0.23 (0.16, 0.33)	<0.001	0
	Sorafenib	1	187	0.1 (0.03, 0.3)	<0.001	0.1 (0.03, 0.3)	<0.001	0
	Bev + IFN	1	641	1.05 (0.88, 1.26)	0.55	1.05 (0.88, 1.26)	0.55	0
	Temsirolimus	1	408	0.48 (0.36, 0.64)	<0.001	0.48 (0.36, 0.64)	<0.001	0
Skin and subcutaneous tissue disorders								
Alopecia	Sunitinib	1	735	1.35 (0.88, 2.07)	0.17	1.35 (0.88, 2.07)	0.17	0
	Sorafenib	1	187	7.42 (3.07, 17.97)	<0.001	7.42 (3.07, 17.97)	<0.001	0
Hair colour change	Sunitinib	1	735	18 (6.65, 48.7)	<0.001	18 (6.65, 48.7)	<0.001	0
Hand-foot syndrome	Sunitinib	1	735	9.51 (5.21, 17.38)	<0.001	9.51 (5.21, 17.38)	<0.001	0
	Sorafenib	1	187	13.45 (5.09, 35.55)	<0.001	13.45 (5.09, 35.55)	<0.001	0
Rash	Sunitinib	1	735	2.98 (2.01, 4.41)	<0.001	2.98 (2.01, 4.41)	<0.001	0
	Sorafenib	1	187	4.64 (2.3, 9.37)	<0.001	4.64 (2.3, 9.37)	<0.001	0
	Temsirolimus	1	408	7.85 (4.45, 13.85)	<0.001	7.85 (4.45, 13.85)	<0.001	0
Skin discolouration	Sunitinib	1	735	24.24 (9.02, 65.15)	<0.001	24.24 (9.02, 65.15)	<0.001	0
Investigations								
ALT increased	Sunitinib	1	735	1.27 (1.08, 1.5)	<0.001	1.27 (1.08, 1.5)	<0.001	0
AST increased	Sunitinib	1	735	1.47 (1.25, 1.73)	<0.001	1.47 (1.25, 1.73)	<0.001	0
	Temsirolimus	1	408	0.58 (0.33, 1.03)	0.06	0.58 (0.33, 1.03)	0.06	0
Total bilirubin increased	Sunitinib	1	735	10.29 (4.81, 22.02)	<0.001	10.29 (4.81, 22.02)	<0.001	0
Vascular disorders								
Hypertension	Sunitinib	1	735	7.75 (4.53, 13.25)	<0.001	7.75 (4.53, 13.25)	<0.001	0
	Sorafenib	1	187	4.08 (1.61, 10.32)	<0.001	4.08 (1.61, 10.32)	<0.001	0
	Bev + IFN	1	641	2.84 (1.91, 4.21)	<0.001	2.84 (1.91, 4.21)	<0.001	0
Metabolism and nutrition disorders								
Anorexia	Sunitinib	1	735	1.22 (0.98, 1.51)	0.08	1.22 (0.98, 1.51)	0.08	0
	Sorafenib	1	187	1 (0.64, 1.55)	0.99	1 (0.64, 1.55)	0.99	0
	Bev + IFN	1	641	1.19 (0.95, 1.48)	0.13	1.19 (0.95, 1.48)	0.13	0
	Temsirolimus	1	408	0.73 (0.57, 0.94)	0.02	0.73 (0.57, 0.94)	0.02	0
Hyperglycaemia	Temsirolimus	1	408	2.36 (1.5, 3.72)	<0.001	2.36 (1.5, 3.72)	<0.001	0

Outcome	Intervention	Studies	N	Fixed effects		Random effects		I2
				RR (95% CI)	P value	RR (95% CI)	P value	
Hypophosphataemia	Sunitinib	1	735	1.29 (1.02, 1.64)	0.03	1.29 (1.02, 1.64)	0.03	0
Musculoskeletal and connective tissue disorders								
Arthralgia	Sunitinib	1	735	0.79 (0.53, 1.16)	0.23	0.79 (0.53, 1.16)	0.23	0
Nervous system disorders								
Altered taste	Sunitinib	1	735	3.08 (2.35, 4.03)	<0.001	3.08 (2.35, 4.03)	<0.001	0
Headache	Sunitinib	1	735	0.88 (0.62, 1.24)	0.45	0.88 (0.62, 1.24)	0.45	0
	Bev + IFN	1	641	1.45 (1.06, 2)	0.02	1.45 (1.06, 2)	0.02	0
	Temsirolimus	1	408	0.99 (0.63, 1.58)	0.98	0.99 (0.63, 1.58)	0.98	0
Respiratory, thoracic and mediastinal disorders								
Epistaxis	Sunitinib	1	735	9.33 (4.34, 20.03)	<0.001	9.33 (4.34, 20.03)	<0.001	0
Infections and infestations								
Infection	Temsirolimus	1	408	1.92 (1.28, 2.9)	<0.001	1.92 (1.28, 2.9)	<0.001	0
Flu-like symptoms	Sorafenib	1	187	0.19 (0.07, 0.52)	<0.001	0.19 (0.07, 0.52)	<0.001	0
	Bev + IFN	1	641	0.96 (0.73, 1.26)	0.77	0.96 (0.73, 1.26)	0.77	0
Blood and lymphatic system disorders								
Anaemia	Sunitinib	1	735	1.13 (1.04, 1.23)	0.01	1.13 (1.04, 1.23)	0.01	0
	Bev + IFN	1	641	0.73 (0.47, 1.12)	0.15	0.73 (0.47, 1.12)	0.15	0
	Temsirolimus	1	408	1.08 (0.86, 1.34)	0.52	1.08 (0.86, 1.34)	0.52	0
Leucopenia	Sunitinib	1	735	1.37 (1.24, 1.52)	<0.001	1.37 (1.24, 1.52)	<0.001	0
	Temsirolimus	1	408	0.34 (0.18, 0.64)	<0.001	0.34 (0.18, 0.64)	<0.001	0
Lymphocytopaenia	Sunitinib	1	735	0.99 (0.89, 1.09)	0.8	0.99 (0.89, 1.09)	0.8	0
Neutropaenia	Sunitinib	1	735	1.54 (1.37, 1.73)	<0.001	1.54 (1.37, 1.73)	<0.001	0
	Bev + IFN	1	641	1.08 (0.61, 1.92)	0.79	1.08 (0.61, 1.92)	0.79	0
	Temsirolimus	1	408	0.6 (0.32, 1.11)	0.1	0.6 (0.32, 1.11)	0.1	0
Thrombocytopenia	Sunitinib	1	735	2.6 (2.16, 3.14)	<0.001	2.6 (2.16, 3.14)	<0.001	0
	Bev + IFN	1	641	1.58 (0.79, 3.15)	0.2	1.58 (0.79, 3.15)	0.2	0
	Temsirolimus	1	408	1.74 (0.98, 3.11)	0.06	1.74 (0.98, 3.11)	0.06	0
Psychiatric disorders								
Depression	Sorafenib	1	187	0.03 (0, 0.57)	0.02	0.03 (0, 0.57)	0.02	0
	Bev + IFN	1	641	1.19 (0.77, 1.85)	0.43	1.19 (0.77, 1.85)	0.43	0
Congestive heart failure	Bev + IFN	1	641	0.9 (0.06, 14.36)	0.94	0.9 (0.06, 14.36)	0.94	0
Endocrine disorders								
Hypothyroidism	Sunitinib	1	735	7.27 (3.35, 15.77)	<0.001	7.27 (3.35, 15.77)	<0.001	0

6.4 Tolerability

Reasons for treatment discontinuation were reported in six studies and were analysed (Table 46). Treatment discontinuation due to AEs was higher in patients receiving bevacizumab plus IFN group compared to IFN alone and the difference was statistically significant in favour of the IFN group. However, heterogeneity in trials assessing this outcome limits the validity of these findings.

Table 46: Result of meta-analysis – treatment discontinuations

Outcome	Intervention	Comparator	Studies	N	Fixed effects		Random effects		I2
					RR (95% CI)	P value	RR (95% CI)	P value	
Withdrawals due to any cause	Pazopanib	Placebo/BSC	1	233	0.87 (0.77, 0.98)	0.022	0.87 (0.77, 0.98)	0.022	0
	Sunitinib	IFN	1	750	0.88 (0.84, 0.91)	<0.001	0.88 (0.84, 0.91)	<0.001	0
	Sorafenib	IFN	1	189	0.89 (0.47, 1.69)	0.72	0.89 (0.47, 1.69)	0.72	0
	Bev + IFN	IFN	2	1381	1.11 (1.05, 1.17)	<0.001	1.36 (0.34, 5.43)	0.67	99
	Temsirolimus	IFN	1	416	1.02 (0.97, 1.06)	0.51	1.02 (0.97, 1.06)	0.51	0
	IFN	Placebo/BSC	1	160	0.81 (0.71, 0.92)	<0.001	0.81 (0.71, 0.92)	<0.001	0
Withdrawals due to AEs	Pazopanib	Placebo/BSC	1	233	2.14 (0.74, 6.14)	0.158	2.14 (0.74, 6.14)	0.158	0
	Sunitinib	IFN	1	750	0.81 (0.61, 1.08)	0.15	0.81 (0.61, 1.08)	0.15	0
	Sorafenib	IFN	1	189	0.75 (0.36, 1.56)	0.43	0.75 (0.36, 1.56)	0.43	0
	Bev + IFN	IFN	2	1381	1.72 (1.38, 2.14)	<0.001	1.78 (0.9, 3.5)	0.1	89

Outcome	Intervention	Comparator	Studies	N	Fixed effects		Random effects		I2
					RR (95% CI)	P value	RR (95% CI)	P value	
	Temsirolimus	IFN	1	416	0.51 (0.28, 0.93)	0.03	0.51 (0.28, 0.93)	0.03	0
	IFN	Placebo/BSC	1	160	11.28 (0.63, 200.58)	0.1	11.28 (0.63, 200.58)	0.1	0
Withdrawals due to death	Pazopanib	Placebo/BSC	1	233	0.5 (0.17, 1.51)	0.22	0.5 (0.17, 1.51)	0.22	0
	Sunitinib	IFN	1	750	1.15 (0.64, 2.06)	0.64	1.15 (0.64, 2.06)	0.64	0
	Sorafenib	IFN	1	189	1.90 (0.36, 10.11)	0.45	1.90 (0.36, 10.11)	0.45	0
	Bev + IFN	IFN	1	649	1.13 (0.41, 3.07)	0.82	1.13 (0.41, 3.07)	0.82	0
	Temsirolimus	IFN	1	416	0.59 (0.22, 1.61)	0.3	0.59 (0.22, 1.61)	0.3	0
	IFN	Placebo/BSC	1	160	0.34 (0.07, 1.64)	0.18	0.34 (0.07, 1.64)	0.18	0

BEV = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha, ITT = intention to treat, RR = relative risk.

6.5 Dose interruptions and dose reductions

The results of meta-analysis indicate that dose interruptions and dose reductions were numerically higher in the intervention group compared to control group. The only exception to this was temsirolimus which showed favourable and statistically significant result compared to IFN for in terms of dose reductions.

Table 47: Result of meta-analysis – down dosing and dose interruption

Outcome	Intervention	Comparator	Studies	N	Fixed effects		Random effects		I2
					RR (95% CI)	P value	RR (95% CI)	P value	
Dose interruption	Pazopanib	Placebo/BSC	1	233	7.17 (2.7, 19.04)	<0.001	7.17 (2.7, 19.04)	<0.001	0
	Sunitinib	IFN	1	750	1.19 (0.98, 1.45)	0.08	1.19 (0.98, 1.45)	0.08	0
	Sorafenib	IFN	1	189	1.4 (1.05, 1.86)	0.02	1.4 (1.05, 1.86)	0.02	0
	Bev + IFN	IFN	1	732	1.95 (1.64, 2.31)	<0.001	1.95 (1.64, 2.31)	<0.001	0
	Temsirolimus	IFN	1	416	1.01 (0.87, 1.16)	0.94	1.01 (0.87, 1.16)	0.94	0
Dose reduction	Pazopanib	Placebo/BSC	1	233	6.04 (1.92, 18.99)	0.002	6.04 (1.92, 18.99)	0.002	0
	Sunitinib	IFN	1	750	1.86 (1.53, 2.26)	<0.001	1.86 (1.53, 2.26)	<0.001	0
	Sorafenib	IFN	1	189	1.26 (0.81, 1.98)	0.3	1.26 (0.81, 1.98)	0.3	0
	Bev + IFN	IFN	2	1381	1.34 (1.2, 1.51)	<0.001	1.35 (1.2, 1.51)	<0.001	0
	Temsirolimus	IFN	1	416	0.61 (0.45, 0.83)	<0.001	0.61 (0.45, 0.83)	<0.001	0
	IFN	Placebo/BSC	1	350	3.37 (1.83, 6.21)	<0.001	3.37 (1.83, 6.21)	<0.001	0

BEV = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha, ITT = intention to treat, RR = relative risk.

7 Indirect comparisons of results from studies in the clinical systematic review – considering MRC RE01 study only

This section reports the result of indirect comparisons between pazopanib and comparators of interest (sunitinib, sorafenib, bevacizumab plus IFN, temsirolimus) using MRC RE01 study only.

7.1 Summary of trials

There were no studies which directly compared pazopanib with a comparator of interest. Therefore, indirect analyses were performed to investigate the comparative efficacy, safety, and tolerability of pazopanib with the identified comparators for this review. The results of direct meta-analyses of trials (comparing one of the study drugs to either another study drug or to a placebo/BSC) were used as inputs into indirect analyses.

The indirect comparisons which could be carried out were limited by the comparisons performed in the trials identified, and by the outcomes reported. In particular, in several cases, outcomes such as the TTR and DOR were reported without accompanying CIs, preventing any analysis of these data.

Indirect analyses should be performed with caution, and should consider the differences in patient population between trials. The patient population in the temsirolimus trial generally had a poorer prognosis than those enrolled in the pazopanib trial. Approximately two thirds of patients in the temsirolimus trial were categorized as MSKCC poor risk as opposed to approximately 5% of patients in the pazopanib trial.

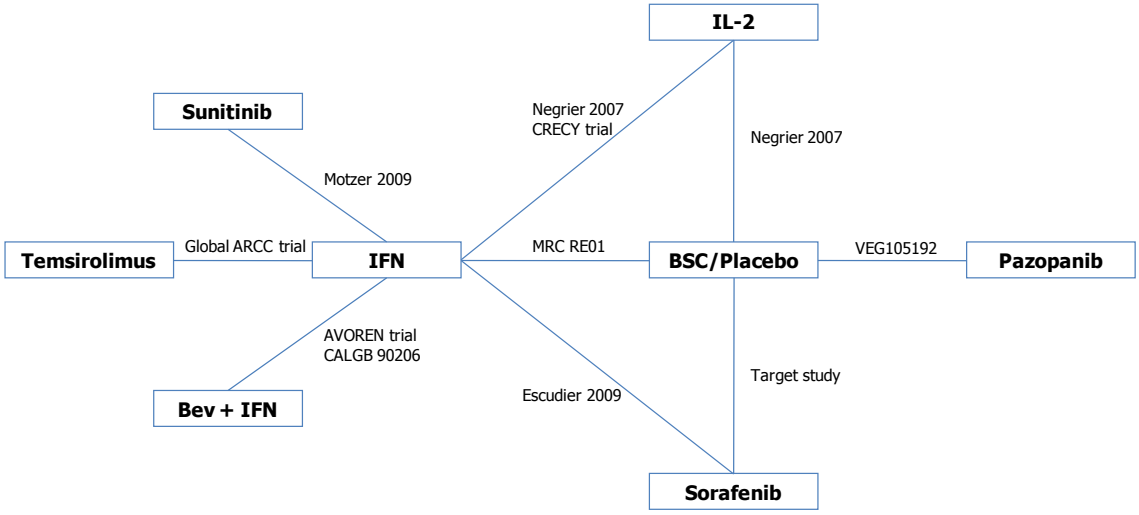
The sunitinib and pazopanib trials appeared comparable; the distribution of patients according to MSKCC risk factor was similar in both trials. In addition, little difference was apparent in the distribution of patients with each number of metastatic site and in the percentage of patients with prior nephrectomy. The sunitinib trial included patients with clear cell histology only while in the pazopanib trial, the majority of patients had clear cell histology (approximately 91%) with the remaining patients having predominantly clear cell histology. The sunitinib study conscripted a higher proportion of patients with ECOG performance status 0 (about 60%) as compared to the pazopanib study where this number was about 40% only. For more details on the differences in patient population between these trials see section 4.4.2 The network diagram showing direct comparison between treatments investigated is presented in Figure 20.

Indirect comparisons of temsirolimus and pazopanib were possible for certain efficacy outcomes, though were not conducted owing to the significantly different patient populations enrolled in the trials contributing to the analyses. Only indirect comparisons between these treatments were conducted for safety outcomes. Further, the temsirolimus trial did not provide a sub-group analysis for patients with less severe disease (MSKCC intermediate risk), thus eliminating the possibility of an indirect comparison with pazopanib. The heterogeneity in patient population between these trials would therefore have been a confounding factor in the results, and the difference obtained from the indirect analysis would not be a true representation of the difference in efficacy, safety and tolerability between the treatments.

For this indirect analysis only one of the five trials, MRC RE01, comparing IFN to BSC was utilised to provide the indirect pathway from pazopanib to the comparators of interest. This study was the only study to provide hazard ratios for OS and PFS, therefore none of the remaining studies could contribute to these analyses.

For response rate, three of the studies (Steineck 1990; Kreigmair 1995; Negrier 2007) were excluded from the analysis since there was a lack of data regarding the baseline characteristics of the included studies, therefore the comparability of these studies to the pazopanib and sunitinib studies could not be determined. Additionally, Pyrhonen 1999 enrolled a greater proportion of patients with an ECOG PS of 1 compared to the pazopanib and sunitinib studies, therefore was excluded from the analysis for response rate.

Figure 20: Network diagram for clinical review



BEV = Bevacizumab, BSC = best supportive care, IFN = interferon alpha, IL-2 = Interleukin 2

7.2 Efficacy outcomes

7.2.1 Progression free survival

7.2.1.1 Comparison with sunitinib

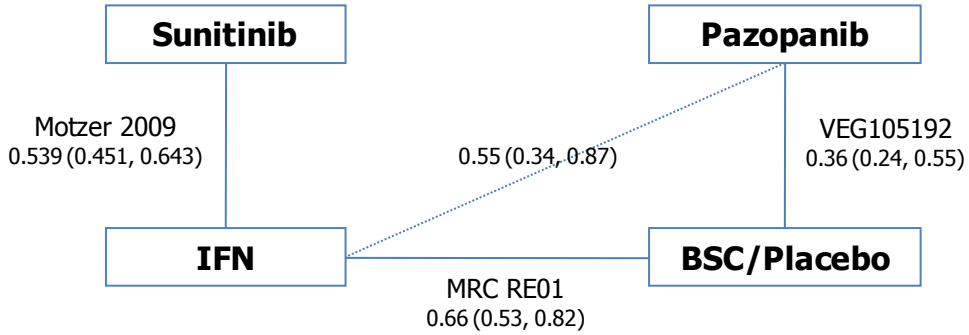
The network diagram for the PFS outcome is shown in Figure 21. This shows the included trials which reported PFS (HR with 95% CI) as an outcome. The indirect comparison of pazopanib with sunitinib was mediated via two common comparators namely, IFN and placebo/BSC.

Indirect comparison between pazopanib and sunitinib involved an intermediate step, comparing IFN to placebo/BSC, using the MRC RE01 trial. All three trials were of similar design with only minor differences in some aspects of the studies (study location, blinding, follow up duration and patient population). The difference between length of follow-up may affect the results of survival analyses. Additionally, the assessment of PFS was performed by IRC in the pazopanib and sunitinib trial however it was not clear whether it was performed by the IRC or investigator in the MRC RE01 trial.

The distribution of patients according to MSKCC risk factor was similar in the pazopanib and sunitinib trial whereas patients were not categorised according to MSKCC score in the MRC RE01 trial. The sunitinib trial included patients with clear cell histology only while in

the pazopanib trial, the majority of patients had clear cell histology (~91%) with the remaining patients having predominantly clear cell histology. The proportion of patients with prior nephrectomy was similar in all three trials. The sunitinib study conscripted higher proportion of patients with ECOG performance status 0 (about 60%) as compared to the pazopanib study where this number was about 40% only, whereas patient population in MRC RE01 has mixed ECOG performance status.

Figure 21: Network diagram of trials providing data for indirect analyses for PFS



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to sunitinib for PFS are detailed below:

- The estimated hazard ratio was 1.01 (95% CI: 0.61, 1.67, p = 0.9629)
- Pazopanib showed comparable efficacy to sunitinib with no statistically significant difference (95% confidence interval includes 1).

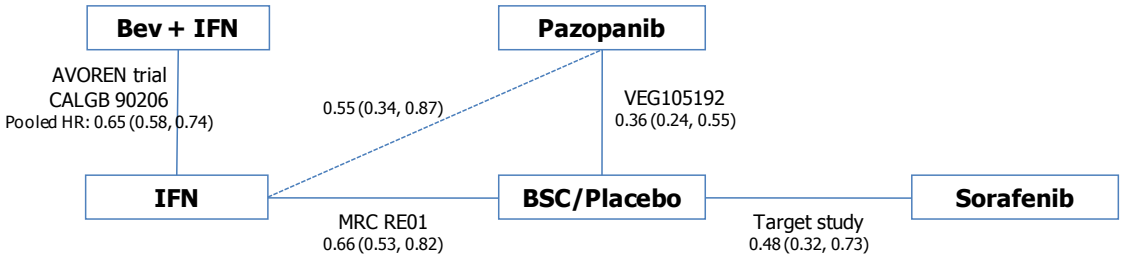
7.2.1.2 Comparison with other comparators

The network diagram for the PFS outcome is shown in Figure 22. This shows the included trials which reported PFS (HR with 95% CI) as an outcome. The indirect comparison with bevacizumab plus IFN was mediated via two common comparators namely, IFN and placebo/BSC, while sorafenib was indirectly compared with pazopanib using placebo as common comparator.

Both sorafenib and pazopanib were investigated in placebo controlled trials, so they could be compared indirectly. Both trials were of similar design and provided subgroup data for treatment-naïve population, although the median follow up duration was lower in TARGET study (28.6 weeks) than the pazopanib trial (58.5 weeks). The definition of PFS was consistent across studies analysed. The assessment of PFS was performed by IRC in the pazopanib trial however it was not clear in the sorafenib trial.

Bevacizumab plus IFN was compared with IFN alone in two studies, namely AVOREN and CALGB 90206 trials. PFS was assessed by investigator in the AVOREN study it was not clear in the CALGB 90206 trial. Both trials were of similar design; however, CALGB 90206 did not include patients from European countries. Results from these studies were pooled and used for the indirect comparison with pazopanib; using MRC RE01 as an intermediate trial. Patient characteristics and study design was comparable across studies used for the indirect comparison of pazopanib and bevacizumab plus IFN.

Figure 22: Network diagram of trials providing data for indirect analyses for PFS



Dotted line shows indirect comparison (intermediate step); BEV = Bevacizumab, BSC = best supportive care, IFN = interferon alpha.

The results from indirect analyses of pazopanib relative to comparators for PFS are shown in Table 48:

- Pazopanib showed a favourable hazard ratio (lower risk) over sorafenib and bevacizumab plus IFN; however this difference did not reach statistical significance.

Table 48: Result of indirect comparison of PFS (pazopanib vs. comparator)

Sorafenib	Bevacizumab plus IFN
0.75 (0.42, 1.35), p = 0.3349	0.84 (0.52, 1.36), p = 0.4778

Reported as Hazard ratio (95% CI)
IFN = Interferon alpha, CI = confidence interval.

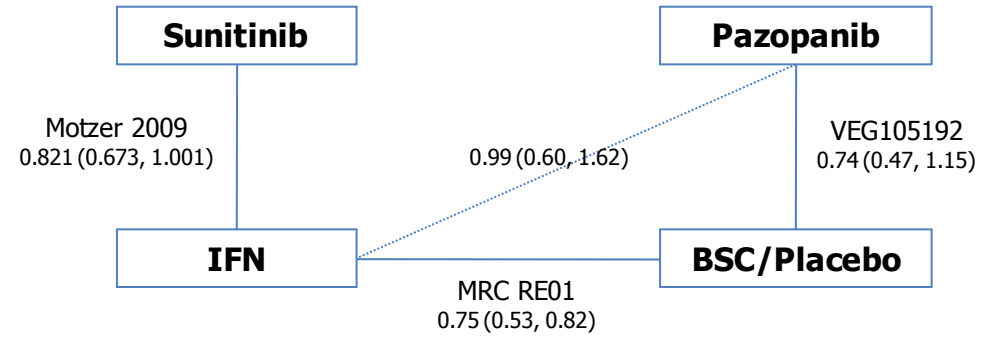
7.2.2 Overall survival

7.2.2.1 Comparison with sunitinib

The network diagram for the outcome OS is shown in Figure 23. This shows the included trials which reported OS suitable for meta-analysis (HR with 95% CI). The definition of OS was similar in all included studies. The pazopanib trial reported the interim OS data in which <50% of patients had died (VEG105192 study) whereas sunitinib trial (Motzer 2000) reported the final data. Hence, indirect comparison of OS should be interpreted with caution due to different levels of maturity of the data in the different trials. The indirect comparison was mediated via two common comparators namely, IFN and placebo/BSC.

Heterogeneity between trials used for the indirect comparison of pazopanib with sunitinib has been discussed in section 7.2.1.1.

Figure 23: Network diagram of trials providing data for indirect analyses for OS



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to sunitinib for OS are detailed

below:

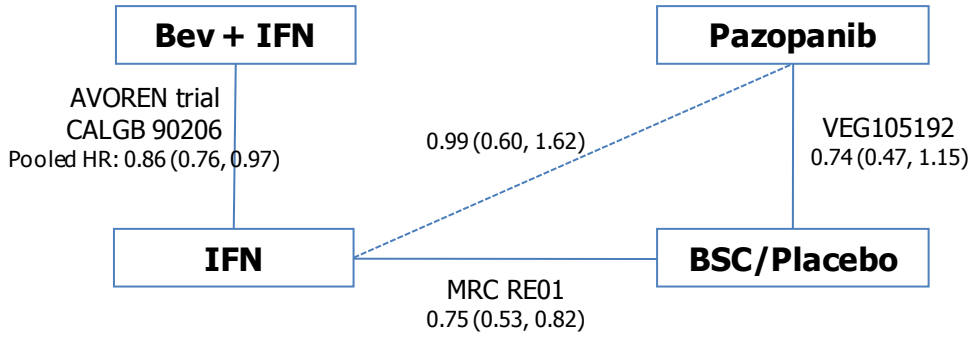
- The estimated hazard ratio was 1.20 (95% CI: 0.70, 2.05, p = 0.5014)
- Pazopanib showed similar efficacy against sunitinib with no statistically significant difference (95% confidence interval include 1).

7.2.2.2 Comparison with other comparators

The network diagram for the outcome OS is shown in Figure 24. This shows the included trials which reported OS suitable for meta-analysis (HR with 95% CI). The indirect comparison was mediated via two common comparators namely, IFN and placebo/BSC.

Heterogeneity between trials used for the indirect comparison of pazopanib with bevacizumab plus IFN has been discussed in section 7.2.1.2. Sorafenib was investigated in two trials though none of these provided a hazard ratio for OS and hence could not be compared with pazopanib indirectly.

Figure 24: Network diagram of trials providing data for indirect analyses for OS



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to comparators for OS are shown in Table 49:

- Pazopanib showed similar efficacy against bevacizumab plus IFN with no statistically significant difference (95% confidence interval include 1).

Table 49: Results from indirect comparison of OS (pazopanib vs. comparator)

Bevacizumab plus IFN	Sorafenib
1.15 (0.69, 1.92), p = 0.5993	-

Reported as Hazard ratio (95% CI)
BEV = Bevacizumab, BSC = best supportive care, IFN = interferon alpha, OS = overall survival.

7.2.3 Response rate

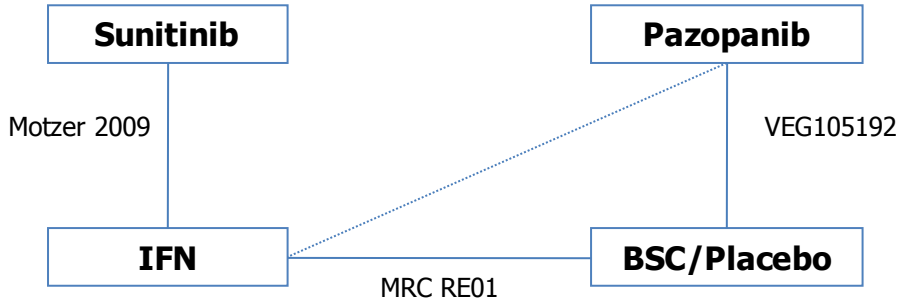
7.2.3.1 Comparison with sunitinib

Response rate was a commonly reported secondary outcome measure. The network diagram for the response rate outcome is shown in Figure 25. This shows the included trials which reported response rate data.

Similar to PFS outcome, the indirect comparison with sunitinib was mediated via two common comparators namely, IFN and placebo/BSC. Heterogeneity between these trials has been discussed in section 7.2.1.1. Data on ORR was taken from the MRC RE01 study for the 6 months endpoint only (rather than the 12 weeks endpoint) as a better marker for response over a 6 month period. Additionally, the assessment of response rate was performed by IRC in the pazopanib and sunitinib trial however it was not clear in the MRC

RE01 trial. Both pazopanib and sunitinib trial reported investigator assessed and IRC assessed response rate.

Figure 25: Network diagram of trials providing data for indirect analyses for response rate (MRC RE01 study only)



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to sunitinib for response rate are shown in Table 50.

- Pazopanib showed a numerically lower ORR, CR, and PR as compared to sunitinib; however, this was not statistically significant.

Table 50: Results from indirect comparison of response rate (pazopanib vs. sunitinib)

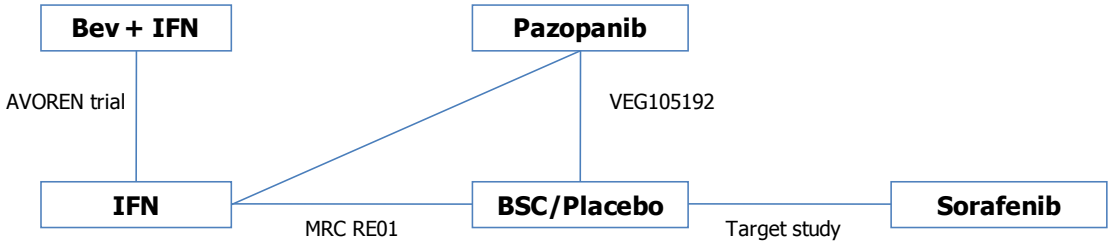
Comparison	Assessed by	Response rate		
		Overall	Complete	Partial
Pazopanib vs. sunitinib	IRC	0.57 (0.11, 3.02), p = 0.5122	-	0.57 (0.11, 3.02), p = 0.5122
Pazopanib vs. sunitinib	Investigator	0.57 (0.13, 2.42), p = 0.4444	0.18 (0, 15.26), p = 0.4509	0.65 (0.15, 2.88), p = 0.573

Reported as risk ratio (95% CI), IRC = Independent review committee

7.2.3.2 Comparison with other comparators

ORR was a commonly reported secondary outcome measure. The network diagram for the ORR outcome is shown in Figure 26. This shows the included trials which reported ORR data. Similar to PFS outcome, the indirect comparison with bevacizumab plus IFN was mediated via two common comparators, namely IFN and placebo/BSC while sorafenib was indirectly compared with pazopanib using placebo as the common comparator. Heterogeneity between these is discussed above in section 7.2.1.2. The assessment of response rate was performed by IRC in both pazopanib and sorafenib trial however it was by investigator in Bev + IFN trial. Data on response rate was taken from the MRC RE01 study for the 6 months endpoint only (rather than the 12 weeks endpoint) as a better marker for response over a 6 month period.

Figure 26: Network diagram of trials providing data for indirect analyses for response rate (MRC RE01 study only)



The result from indirect analysis of pazopanib relative to sorafenib for ORR is detailed in Table 51.

- Pazopanib showed a numerically lower, but not statistically significantly different response rate to sorafenib (95% confidence intervals include 1).
- Pazopanib showed a numerically lower, but not statistically significantly different response rate to Bev + IFN (95% confidence intervals include 1).

Table 51: Results from indirect comparison of response rate (pazopanib vs. comparators)

Comparison	Assessed by	Response rate		
		Overall	Complete	Partial
Pazopanib vs. sorafenib	IRC	0.44 (0.02, 9.4), p = 0.602	-	0.5 (0.02, 10.76), p = 0.6601
Pazopanib vs. sorafenib	Investigator	-	-	-
Pazopanib vs. Bev + IFN	IRC	-	-	-
Pazopanib vs. Bev + IFN	Investigator	0.85 (0.2, 3.67), p = 0.8274	0.76 (0.01, 66.02), p = 0.9053	0.88 (0.2, 3.93), p = 0.8645

Reported as risk ratio (95% CI), IRC = Independent review committee

7.3 AEs

The MRC RE01 trial did not report data for all grade AEs hence indirect comparison of pazopanib with sunitinib and bevacizumab plus IFN was not possible. The indirect comparison between pazopanib and sorafenib is shown in section 8.3.2.

8 Indirect comparisons of results from studies in the clinical systematic review – considering all IFN studies

This section reports the result of indirect comparisons between pazopanib and comparator of interest (sunitinib, sorafenib, bevacizumab plus IFN, temsirolimus) using all IFN studies. We have considered all studies (eliminating selection bias) for the purpose of indirect comparison however, this do not mean that all studies has contributed in the analysis. Availability of reported outcomes may limit number of studies used in actual analysis.

8.1 Summary of trials

There were no studies which directly compared pazopanib with a comparator of interest. Therefore, indirect analyses were performed to investigate the comparative efficacy, safety, and tolerability of pazopanib with the identified comparators for this review. The results of direct meta-analyses of trials (comparing one of the study drugs to either another study drug or to a placebo/BSC) were used as inputs into indirect analyses.

The indirect comparisons which could be carried out were limited by the comparisons performed in the trials identified, and by the outcomes reported. In particular, in several cases, outcomes such as the TTR and DOR were reported without accompanying CIs, preventing any analysis of these data.

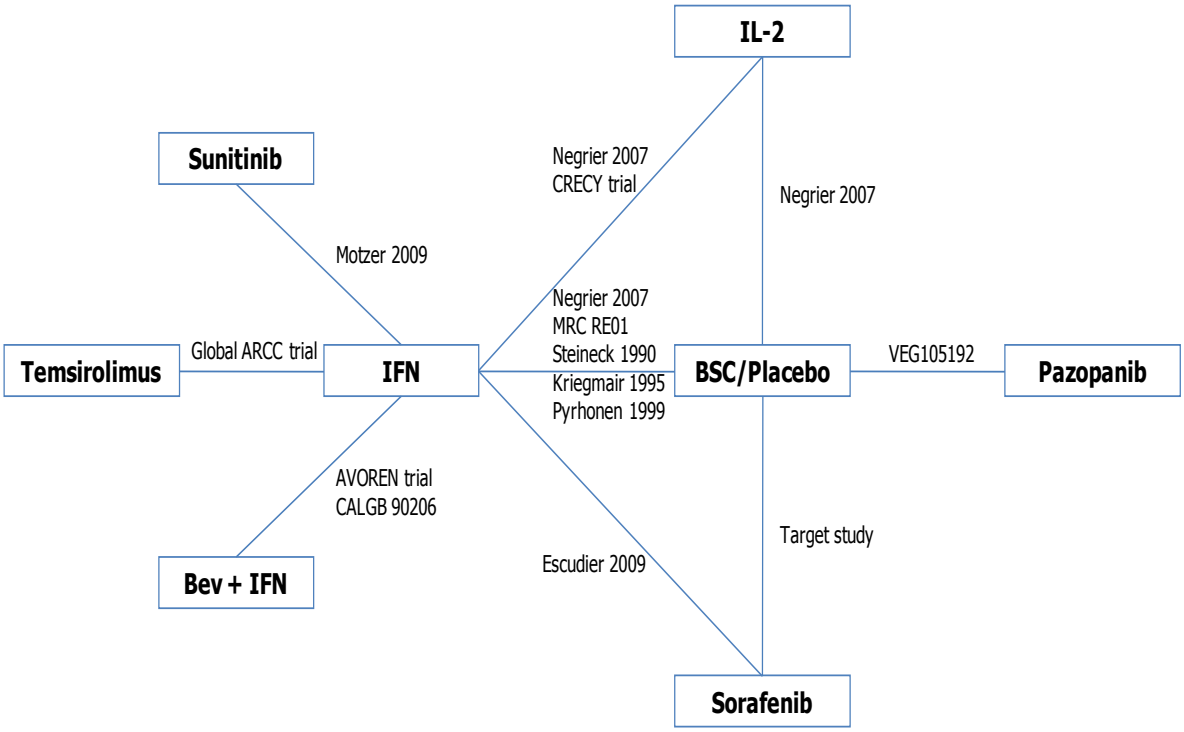
Indirect analyses should be performed with caution, and should consider the differences in patient population between trials. The patient population in the temsirolimus trial generally had a poorer prognosis than those enrolled in the pazopanib trial. Approximately two thirds of patients in the temsirolimus trial were categorized as MSKCC poor risk as opposed to approximately 5% of patients in the pazopanib trial.

The sunitinib and pazopanib trials appeared comparable; the distribution of patients according to MSKCC risk factor similar in both trials. In addition, little difference was apparent in the distribution of patients with each number of metastatic site and in the percentage of patients with prior nephrectomy. The sunitinib trial included patients with clear cell histology only while in the pazopanib trial, the majority of patients had clear cell histology (approximately 91%) with the remaining patients having predominantly clear cell histology. The sunitinib study conscripted higher proportion of patients with ECOG performance status 0 (about 60%) as compared to the pazopanib study where this number was about 40% only. For more details on the differences in patient population between these trials see section 4.4.2. The network diagram showing direct comparison between treatments investigated is presented in Figure 27.

Indirect comparisons of temsirolimus and pazopanib were possible for certain efficacy outcomes, though were not conducted owing to the significantly different patient populations enrolled in the trials contributing to the analyses. Only indirect comparisons between these treatments were conducted for safety outcomes. Further, the temsirolimus trial did not provide a sub-group analysis for patients with less severe disease (MSKCC intermediate risk), thus eliminating the possibility of an indirect comparison with pazopanib. The heterogeneity in patient population between these trials

would therefore have been a confounding factor in the results, and the difference obtained from the indirect analysis would not be a true representation of the difference in efficacy, safety and tolerability between the treatments.

Figure 27: Network diagram for clinical review



BEV = Bevacizumab, BSC = best supportive care, CI = confidence interval, IFN = interferon alpha, IL-2 = Interleukin 2

8.2 Efficacy outcomes

8.2.1 Progression free survival

8.2.1.1 Comparison with sunitinib

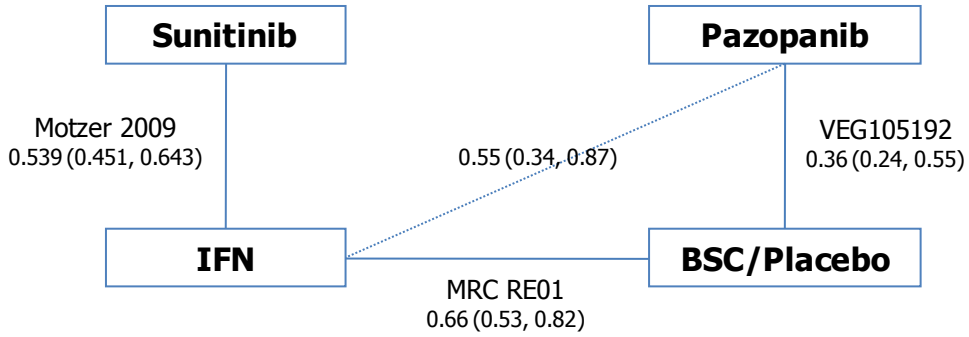
The network diagram for the PFS outcome is shown in Figure 28. This shows the included trials which reported PFS (HR with 95% CI) as an outcome. The indirect comparison of pazopanib with sunitinib was mediated via two common comparators namely, IFN and placebo/BSC.

Indirect comparison between pazopanib and sunitinib involved an intermediate step, comparing IFN to placebo/BSC, using the MRC RE01 trial. All three trials were of similar design with only minor differences in some aspects of the studies (study location, blinding, follow up duration and patient population). The difference between length of follow-up may affect the results of survival analyses. Additionally, the assessment of PFS was performed by IRC in the pazopanib and sunitinib trial however it was not clear in the MRC RE01 trial.

The distribution of patients according to MSKCC risk factor was similar in the pazopanib and sunitinib trial whereas it was not reported for the MRC RE01 trial. The sunitinib trial included patients with clear cell histology only while in the pazopanib trial, the majority of patients had clear cell histology (~91%) with the remaining patients having predominantly clear cell histology. The proportion of patients with prior nephrectomy was

similar in all three trials. The pazopanib study conscripted higher proportion of patients with ECOG performance status1 (about 60%) as compared to the sunitinib study where this number was about 40% only, whereas patient population in MRC RE01 has mixed ECOG performance status.

Figure 28: Network diagram of trials providing data for indirect analyses for PFS



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to sunitinib for PFS are detailed below:

- The estimated hazard ratio was 1.01 (95% CI: 0.61, 1.67, p = 0.9629)
- Pazopanib showed comparable efficacy to sunitinib with no statistically significant difference (95% confidence interval includes 1).

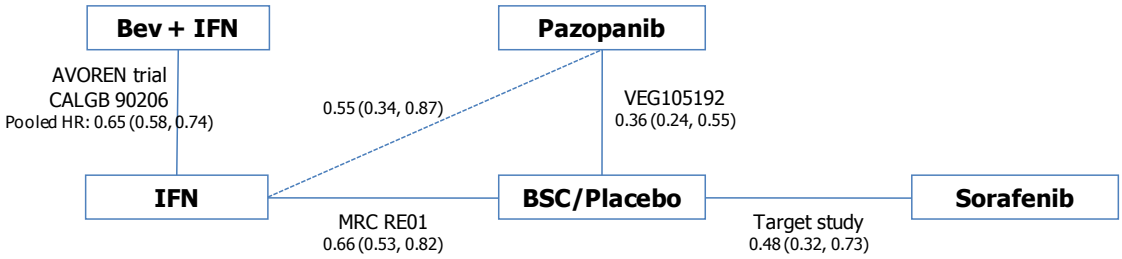
8.2.1.2 Comparison with other comparators

The network diagram for the PFS outcome is shown in Figure 29. This shows the included trials which reported PFS (HR with 95% CI) as an outcome. The indirect comparison with bevacizumab plus IFN was mediated via two common comparators namely, IFN and placebo/BSC, while sorafenib was indirectly compared with pazopanib using placebo as common comparator.

Both sorafenib and pazopanib were investigated in placebo controlled trials, so they could be compared indirectly. Both trials were of similar design and provided subgroup data for the treatment-naïve population, although the median follow up duration was lower in TARGET study (28.6 weeks) than the pazopanib trial (58.5 weeks). The definition of PFS was consistent across studies analysed. The assessment of PFS was performed by the IRC in the pazopanib trial however it was not clear in the sorafenib trial whether this was IRC or investigator reported.

Bevacizumab plus IFN was compared with IFN alone in two studies, namely AVOREN and CALGB 90206 trials. PFS was assessed by investigator in the AVOREN study it was not clear in the CALGB 90206 trial. Both trials were of similar design; however, CALGB 90206 did not include patients from European countries. Results from these studies were pooled and used for the indirect comparison with pazopanib; using MRC RE01 as an intermediate trial. Patient characteristics and study design was comparable across studies used for the indirect comparison of pazopanib and bevacizumab plus IFN.

Figure 29: Network diagram of trials providing data for indirect analyses for PFS



Dotted line shows indirect comparison (intermediate step); BEV = Bevacizumab, BSC = best supportive care, IFN = interferon alpha.

The results from indirect analyses of pazopanib relative to comparators for PFS are shown in Table 52:

- Pazopanib showed a favourable hazard ratio (lower risk) over sorafenib and bevacizumab plus IFN; however this difference did not reach statistical significance.

Table 52: Result of indirect comparison of PFS (pazopanib vs. comparator)

Sorafenib	Bevacizumab plus IFN
0.75 (0.42, 1.35), p = 0.3349	0.84 (0.52, 1.36), p = 0.4778

Reported as Hazard ratio (95% CI)
IFN = Interferon alpha, CI = confidence interval.

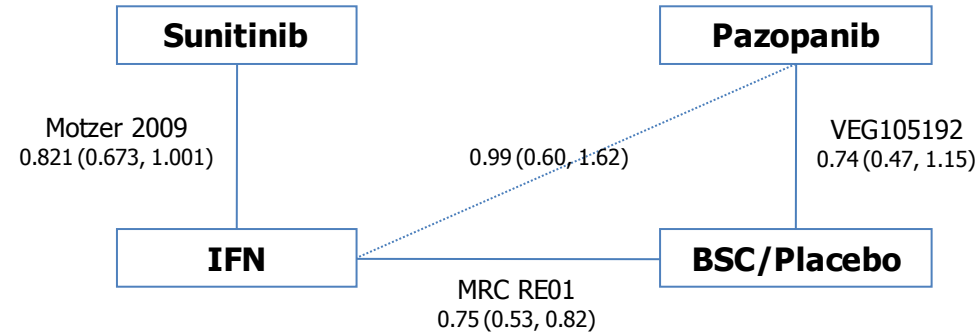
8.2.2 Overall survival

8.2.2.1 Comparison with sunitinib

The network diagram for the outcome OS is shown in Figure 30. This shows the included trials which reported OS suitable for meta-analysis (HR with 95% CI). The definition of OS was similar in all included studies. The pazopanib trial reported the interim OS data in which <50% of patients had died (VEG105192 study) whereas sunitinib trial (Motzer 200) reported the final data. Hence, indirect comparison of OS should be interpreted with caution due to different levels of maturity of the data in the different trials. The indirect comparison was mediated via two common comparators namely, IFN and placebo/BSC.

Heterogeneity between trials used for the indirect comparison of pazopanib with sunitinib has been discussed in section 7.2.1.1.

Figure 30: Network diagram of trials providing data for indirect analyses for OS



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to sunitinib for OS are detailed

below:

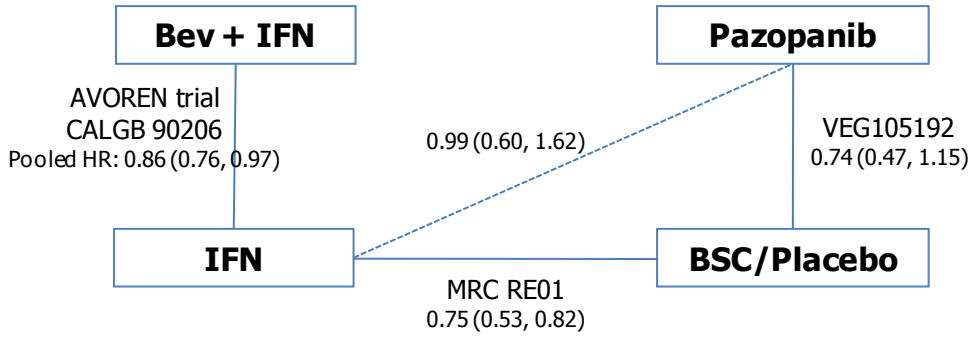
- The estimated hazard ratio was 1.20 (95% CI: 0.70, 2.05, p = 0.5014)
- Pazopanib showed similar efficacy against sunitinib with no statistically significant difference (95% confidence interval include 1).

8.2.2.2 Comparison with other comparators

The network diagram for the outcome OS is shown in Figure 31. This shows the included trials which reported OS suitable for meta-analysis (HR with 95% CI). The indirect comparison was mediated via two common comparators namely, IFN and placebo/BSC.

Heterogeneity between trials used for the indirect comparison of pazopanib with bevacizumab plus IFN has been discussed in section 7.2.1.2. Sorafenib was investigated in two trials though none of these provided a hazard ratio for OS and hence could not be compared with pazopanib indirectly.

Figure 31: Network diagram of trials providing data for indirect analyses for OS



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to comparators for OS are shown in Table 53.

- Pazopanib showed similar efficacy against bevacizumab plus IFN with no statistically significant difference (95% confidence interval include 1).

Table 53: Results from indirect comparison of OS (pazopanib vs. comparator)

Bevacizumab plus IFN	Sorafenib
1.15 (0.69, 1.92), p = 0.5993	-

Reported as Hazard ratio (95% CI)
BEV = Bevacizumab, BSC = best supportive care, IFN = interferon alpha, OS = overall survival.

8.2.3 Overall response rate

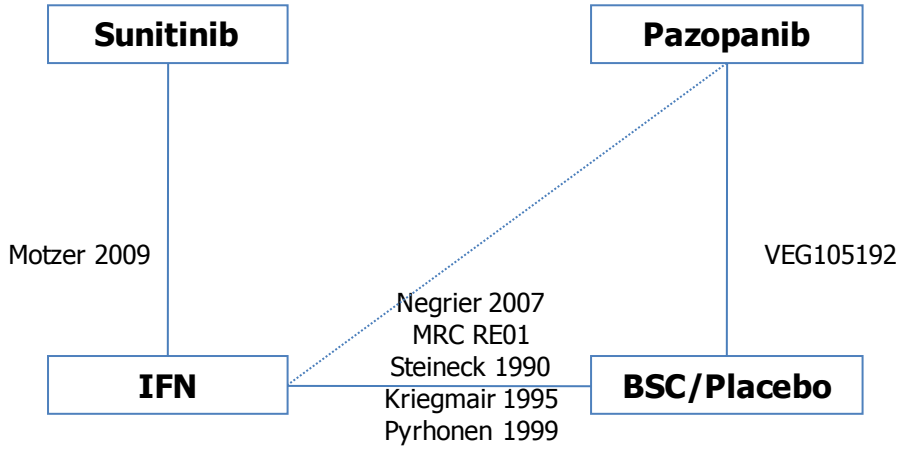
8.2.3.1 Comparison with sunitinib

Response rates were a commonly reported secondary outcome measure. The network diagram for the response rate outcome is shown in Figure 32. This shows the included trials which reported response rate data.

Similar to PFS outcome, the indirect comparison with sunitinib was mediated via two common comparators namely, IFN and placebo/BSC. Heterogeneity between these trials has been discussed in section 7.2.1.1. Data on ORR was taken from the MRC RE01 study for the 6 months endpoint only (rather than the 12 weeks endpoint) as a better marker for response over a 6 month period. Additionally, the assessment of response rate was performed by IRC in the pazopanib and sunitinib trial however it was not clear in the MRC

RE01 trial. Both pazopanib and sunitinib trial reported investigator assessed and IRC assessed response rate.

Figure 32: Network diagram of trials providing data for indirect analyses for response rate (MRC RE01 study only)



Dotted line shows indirect comparison (intermediate step); BSC = best supportive care, IFN = interferon alpha

The results from indirect analyses of pazopanib relative to sunitinib for response rates are shown in Table 54.

- Pazopanib showed a numerically lower ORR, CR and PR as compared to sunitinib; however, this was not statistically significant (95% confidence intervals includes one).

Table 54: Results from indirect comparison of response rate (pazopanib vs. sunitinib)

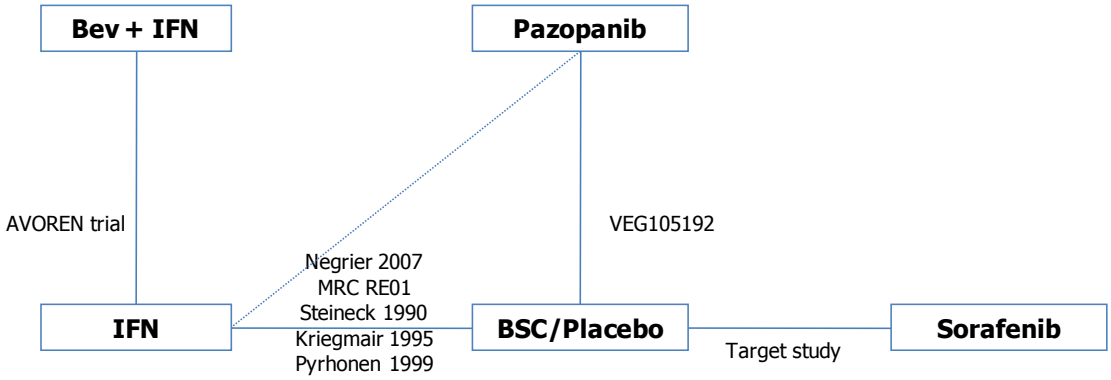
Comparison	Assessed by	Response rate		
		Overall	Complete	Partial
Pazopanib vs. sunitinib	IRC	0.37 (0.09, 1.53), p = 0.1717	-	0.37 (0.09, 1.53), p = 0.1717
Pazopanib vs. sunitinib	Investigator	0.37 (0.12, 1.18), p = 0.092	0.22 (0.01, 6.87), p = 0.3921	0.4 (0.11, 1.38), p = 0.1463

Reported as risk ratio (95% CI), IRC = Independent review committee

8.2.3.2 Comparison with other comparators

Response rates were a commonly reported secondary outcome measure. The network diagram for the response rate outcome is shown in Figure 33. This shows the included trials which reported response rate data. Similar to PFS outcome, the indirect comparison with bevacizumab plus IFN was mediated via two common comparators, namely IFN and placebo/BSC while sorafenib was indirectly compared with pazopanib using placebo as the common comparator. Heterogeneity between these is discussed above in section 7.2.1.2. The assessment of response rate was performed by IRC in both pazopanib and sorafenib trial however it was investigator reported in bevacizumab plus IFN trial. Data on response rate was taken from the MRC RE01 study for the 6 months endpoint only (rather than the 12 weeks endpoint) as a better marker for response over a 6 month period.

Figure 33: Network diagram of trials providing data for indirect analyses for response rate (MRC RE01 study only)



The result from indirect analysis of pazopanib relative to sorafenib for response rate is detailed in Table 55:

- Pazopanib showed a numerically lower, but not statistically significantly different ORR and PR to sorafenib and bevacizumab plus IFN (95% confidence intervals include 1).

Table 55: Results from indirect comparison of response rate (pazopanib vs. comparators)

Comparison	Assessed by	Response rate		
		Overall	Complete	Partial
Pazopanib vs. sorafenib	IRC	0.44 (0.02, 9.4), p = 0.602	-	0.5 (0.02, 10.76), p = 0.6601
Pazopanib vs. sorafenib	Investigator	-	-	-
Pazopanib vs. Bev + IFN	IRC	-	-	-
Pazopanib vs. Bev + IFN	Investigator	0.55 (0.17, 1.79), p = 0.3232	0.94 (0.03, 29.99), p = 0.9726	0.53 (0.15, 1.89), p = 0.3305

Reported as risk ratio (95% CI), IRC = Independent review committee

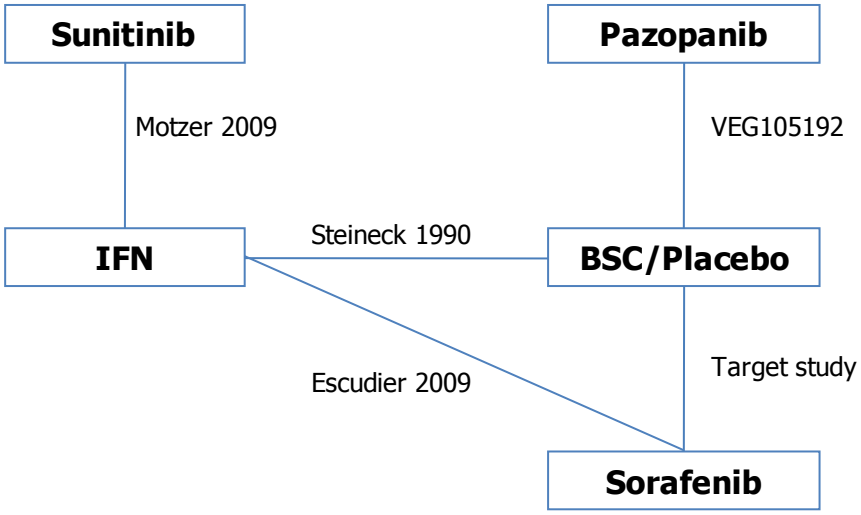
8.3 AEs

8.3.1 Comparison with sunitinib

The network diagram for the AEs is presented in Figure 34. This shows the included trials that reported the percentage of patients who experienced any AE (of all grades).

The comparison of pazopanib with sunitinib was possible via two routes. The first route involves two common comparators, namely placebo/BSC and IFN, and uses data from the Steineck 1990 trial. The Steineck 1990 trial included a total of 60 patients and was conducted in Sweden. Owing to the small trial size, the estimate of treatment effect is unlikely to be precise, hence comparison via an alternative route was considered. The alternative route involves three common comparators, namely placebo/BSC, sorafenib and IFN, and uses data from the TARGET study and the Escudier 2009 study. As the alternative route involves an additional step it is likely to introduce additional heterogeneity. The results have been presented for both of the analyses, where possible.

Figure 34: Network diagram of trials providing data for indirect analyses for AEs



8.3.1.1 Any AEs

Motzer 2009, did not report data for any AEs and therefore cannot be compared with pazopanib indirectly (Motzer 2009).

8.3.1.2 Specific AEs (all grades)

The results from the indirect analyses of pazopanib relative to sunitinib for specific AEs (all grades) are shown in Table 56 and Table 57. Indirect comparison with sunitinib was performed using both possible routes. As expected, the results were not consistent for both routes in many instances.

Pazopanib was associated with a reduced risk (via one or both of the indirect comparison routes) of almost all specific AEs compared to sunitinib where comparison was possible, these included the following:

- Diarrhoea, vomiting, nausea
- Fatigue
- Hand-foot syndrome, rash
- Total bilirubin increased
- Anorexia
- Epistaxis
- Hypertension
- Anaemia, leucopenia, thrombocytopenia

The difference observed, however, rarely reached statistical significance. The reduced risk of the AEs was only observed to be statistically significant for fatigue and diarrhoea via one of the two indirect comparison routes; although both results were contradicted by the result of the alternative route. Alopecia was the only AE where pazopanib showed a consistently (via both routes of indirect comparison) increased but statistically insignificant risk compared to sunitinib. Steineck 1990 did not report data for increased AST & ALT, thus indirect comparison between pazopanib vs. sunitinib was not possible for these outcomes.

Table 56: Result of indirect comparison of AEs (pazopanib vs. sunitinib, via Steineck 1990)

Class	Outcome	Sunitinib
Gastrointestinal disorders	Diarrhoea	0.60 (0.02, 16.11), p = 0.7619
	Vomiting	-
	Nausea	0.56 (0.02, 14.32), p = 0.7289
	Mucositis/stomatitis	-
General disorders and administration site conditions	Fatigue	0.21 (0.06, 0.77), p = 0.0181
Skin and subcutaneous tissue disorders	Alopecia	3.63 (0.05, 253.99), p = 0.5524
	Hand-foot syndrome	-
	Rash	0.23 (0.02, 2.91), p = 0.2535
Investigations	Total bilirubin increased	0.05 (0, 2.55), p = 0.1346
Vascular disorders	Hypertension	2.69 (0.11, 63.56), p = 0.5387
Metabolism and nutrition disorders	Anorexia	0.4 (0.13, 1.29), p = 0.1258
Respiratory, thoracic and mediastinal disorders	Epistaxis	0.09 (0, 7.68), p = 0.2891
Infections and infestations	Flu-like symptoms	-
Blood and lymphatic system disorders	Anaemia	0.28 (0.07, 1.08), p = 0.0652
	Leucopenia	0.14 (0, 7.66), p = 0.3356
	Thrombocytopenia	0.46 (0.01, 17.29), p = 0.6773

Reported as risk ratio (95% CI). Black = point estimate favour pazopanib group; Red = point estimate favour comparator group. CI = confidence interval, IFN = interferon alpha.

Table 57: Result of indirect comparison of AEs (pazopanib vs. sunitinib via Escudier 2009)

Class	Outcome	Sunitinib
Gastrointestinal disorders	Diarrhoea	0.16 (0.04, 0.58), p = 0.0055
	Vomiting	0.36 (0.03, 4.39), p = 0.4198
	Nausea	1.26 (0.38, 4.15), p = 0.7048
General disorders and administration site conditions	Fatigue	0.75 (0.27, 2.07), p = 0.5801
Skin and subcutaneous tissue disorders	Alopecia	9.77 (0.4, 237.93), p = 0.1617
	Hand-foot syndrome	0.82 (0.03, 24.11), p = 0.9088
	Rash	1.21 (0.24, 6.12), p = 0.8163
Vascular disorders	Hypertension	0.2 (0.01, 4.59), p = 0.3103
Metabolism and nutrition disorders	Anorexia	1.63 (0.45, 5.96), p = 0.459

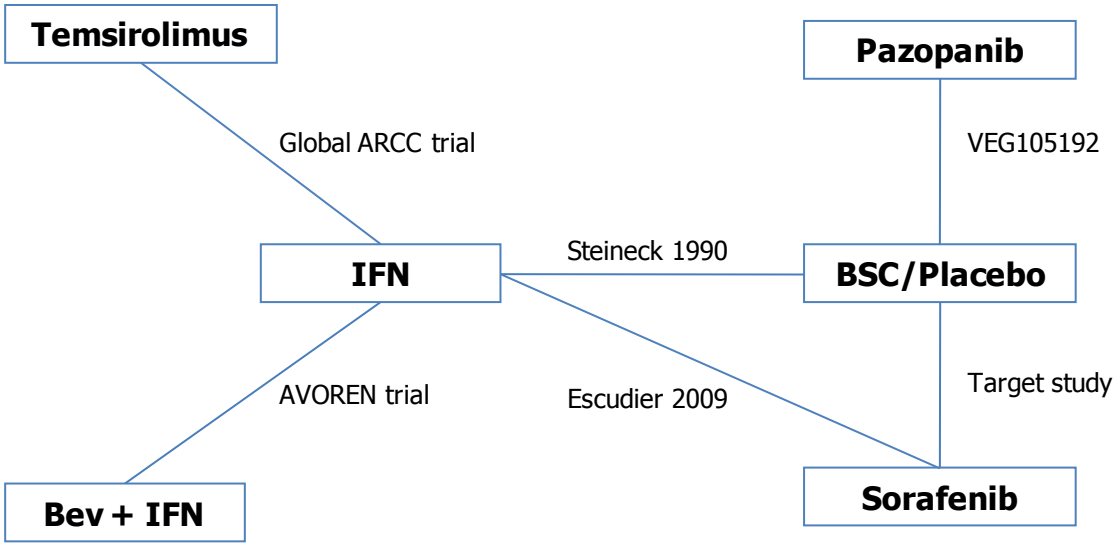
Reported as risk ratio (95% CI). Black = point estimate favour pazopanib group; Red = point estimate favour comparator group. CI = confidence interval, IFN = interferon alpha.

8.3.2 Comparison with other comparators

The network diagram for the AEs is presented in Figure 34. This shows the included trials that reported the percentage of patients who experienced any AE (of all grades). The comparison of pazopanib with sorafenib was mediated via a single common comparator, placebo.

The comparison of pazopanib with bevacizumab plus IFN and temsirolimus was possible via two routes. The first route involves two common comparators, namely placebo/BSC and IFN, and uses data from the Steineck 1990 trial. The Steineck 1990 trial included a total of 60 patients and was conducted in Sweden. Owing to the small trial size, the estimate of treatment effect is unlikely to be precise, hence comparison via an alternative route was considered. The alternative route involves three common comparators, namely placebo/BSC, sorafenib and IFN, and uses data from the TARGET study and the Escudier 2009 study. As the alternative route involves an additional step it is likely to introduce additional heterogeneity. The results have been presented for both of the analyses, where possible.

Figure 35: Network diagram of trials providing data for indirect analyses for AEs



8.3.2.1 Any AEs

Pazopanib showed improved safety (lower risk of any AE/any grade 3 or 4 AE) over sorafenib, bevacizumab plus IFN and temsirolimus. However, the difference was not statistically significant (Table 58 and Table 59).

It should be noted that when pazopanib was compared with bevacizumab plus IFN via the comparison route containing Steineck 1990, the risk of any grade 3 or 4 AE was halved in the pazopanib group compared to bevacizumab plus IFN, a difference which was statistically significant (p = 0.0179). The results were inconsistent as the alternative route showed statistically no significant difference.

Table 58: Result of indirect comparison of AEs (pazopanib vs. comparator, via Steineck 1990 for temsirolimus and bevacizumab plus IFN)

Outcome	Sorafenib	Temsirolimus	Bevacizumab plus IFN
Any AE (all grades)	0.81 (0.6, 1.09), p = 0.1653	-	-
Any grade 3 or 4 AE	0.41 (0.11, 1.50), p = 0.1787	0.59 (0.28, 1.27), p = 0.1792	0.4 (0.19, 0.85), p = 0.0179

Reported as risk ratio (95% CI). AE= adverse event, IFN = interferon alpha.

Table 59: Result of indirect comparison of AEs (pazopanib vs. comparator, via Escudier 2009 for temsirolimus and bevacizumab plus IFN)

Outcome	Sorafenib	Temsirolimus	Bevacizumab plus IFN
Any AE (all grades)	0.81 (0.6, 1.09), p = 0.1653	-	0.84 (0.62, 1.14), p = 0.268
Any grade 3 or 4 AE	0.41 (0.11, 1.50), p = 0.1787	0.56 (0.14, 2.15), p = 0.3952	0.37 (0.1, 1.42), p = 0.1472

Reported as risk ratio (95% CI). AE= adverse event, IFN = interferon alpha.

8.3.2.2 Specific AEs (all grades)

The results from the indirect analyses of pazopanib relative to comparators for specific AEs (all grades) are shown in Table 60 and Table 61. Indirect comparison with bevacizumab plus IFN, and temsirolimus was performed using both possible routes. As expected, the results were not consistent for both routes in many instances.

Pazopanib was associated with a reduced risk of certain AEs compared to sorafenib,

namely vomiting, mucositis/stomatitis, fatigue, hand-foot syndrome, rash, hypertension, flu like symptoms, and anaemia. However, the difference did not reach statistical significance. Compared with sorafenib, pazopanib did not show a statistically significant increase in the risk of the following AEs: diarrhoea, nausea, alopecia, and anorexia.

Compared to bevacizumab plus IFN, pazopanib was associated with a reduced risk of certain AEs namely, fatigue, flu-like symptoms and haematological AEs. The differences did not reach statistical significance, except for the flu-like symptoms outcome.

Results of the temsirolimus versus pazopanib indirect comparisons are inconsistent between the two routes for indirect analysis. Using one route for the indirect analysis, via Steineck 1990, pazopanib demonstrated a reduced risk of nausea, rash, anorexia, leucopenia and thrombocytopenia compared to temsirolimus, with none of the differences reaching significance. Via the alternative route, the risk of nausea and anorexia was increased with pazopanib, although the risk of rash remained lower. Additionally, the increased risk of diarrhoea with pazopanib compared to temsirolimus was not statistically significant via the route containing Steineck 1990; however, it did not reach significance via the alternative route. With these inconsistencies in results, the evidence of comparative safety between temsirolimus and pazopanib is not conclusive.

Across all the treatments pazopanib demonstrated comparable (and for certain AEs, improved) safety profile. This was particularly evident in blood and lymphatic disorder AEs such as anaemia, where the risk of these events was consistently lower with pazopanib than the comparators, although these differences did not reach statistical significance.

Table 60: Result of indirect comparison of AEs (pazopanib vs. comparator, via Steineck 1990 for temsirolimus and bevacizumab plus IFN)

Class	Outcome	Sorafenib	Temsirolimus	Bevacizumab plus IFN
Gastrointestinal disorders	Diarrhoea	2.92 (0.95, 8.98), p = 0.0613	1.82 (0.07, 49.18), p = 0.722	1.85 (0.07, 49.88), p = 0.7146
	Vomiting	0.99 (0.09, 10.81), p = 0.9948	-	-
	Nausea	2.80 (0.98, 8.03), p = 0.0554	0.93 (0.04, 23.67), p = 0.9644	-
	Mucositis/stomatitis	0.57 (0.01, 34.43), p = 0.7908	-	-
General disorders and administration site conditions	Fatigue	0.78 (0.30, 2.02), p= 0.6144	-	0.18 (0.05, 0.68), p = 0.0110
Skin and subcutaneous tissue disorders	Alopecia	1.78 (0.09, 37.05), p = 0.7106	-	-
	Hand-foot syndrome	0.58 (0.02, 13.97), p = 0.7374	-	-
	Rash	0.78 (0.19, 3.17), p = 0.7263	0.09 (0.01, 1.14), p = 0.0628	-
Investigations	Total bilirubin increased	-	-	-
Vascular disorders	Hypertension	0.37 (0.02, 7.22), p = 0.5120	-	7.37 (0.32, 170.15), p = 0.2126
Metabolism and nutrition disorders	Anorexia	1.99 (0.6, 6.60), p = 0.2601	0.56 (0.06, 4.98), p = 0.6019	0.41 (0.13, 1.32), p = 0.1356
Respiratory, thoracic and mediastinal disorders	Epistaxis	-	-	-
Infections and infestations	Flu-like symptoms	0.27 (0.02, 3.29), p = 0.3027	-	0.05 (0.01, 0.52), p = 0.0117
Blood and lymphatic system disorders	Anaemia	0.13 (0.0, 3.85), p = 0.2376	0.29 (0.07, 1.15), p = 0.0791	0.43 (0.11, 1.8), p = 0.2494
	Leucopenia	-	0.57 (0.01, 32.48), p = 0.783	-

Class	Outcome	Sorafenib	Temsirolimus	Bevacizumab plus IFN
	Thrombocytopenia	-	0.69 (0.02, 26.93), p = 0.8443	0.77 (0.02, 30.31), p = 0.8866

Reported as risk ratio (95% CI). Black = point estimate favour pazopanib group; Red = point estimate favour comparator group. CI = confidence interval, IFN = interferon alpha.

Table 61: Result of indirect comparison of AEs (pazopanib vs. comparator, pazopanib vs. comparator, via Escudier 2009 for temsirolimus and bevacizumab plus IFN)

Class	Outcome	Temsirolimus	Bevacizumab plus IFN
Gastrointestinal disorders	Diarrhoea	9.7 (2.61, 36.1), p = 0.0007	0.49 (0.13, 1.83), p = 0.2901
	Vomiting	1.34 (0.11, 16.62), p = 0.8197	-
	Nausea	2.07 (0.62, 6.92), p = 0.2362	-
General disorders and administration site conditions	Fatigue	-	0.66 (0.23, 1.84), p = 0.4213
Skin and subcutaneous tissue disorders	Alopecia	-	-
	Hand-foot syndrome	-	-
	Rash	0.46 (0.09, 2.44), p = 0.3618	-
Vascular disorders	Hypertension	-	0.53 (0.02, 12.28), p = 0.6943
Metabolism and nutrition disorders	Anorexia	2.71 (0.74, 9.96), p = 0.1331	1.28 (0.31, 5.21), p = 0.732

Reported as risk ratio (95% CI). Black = point estimate favour pazopanib group; Red = point estimate favour comparator group. CI = confidence interval, IFN = interferon alpha.

9 Systematic review of economic studies

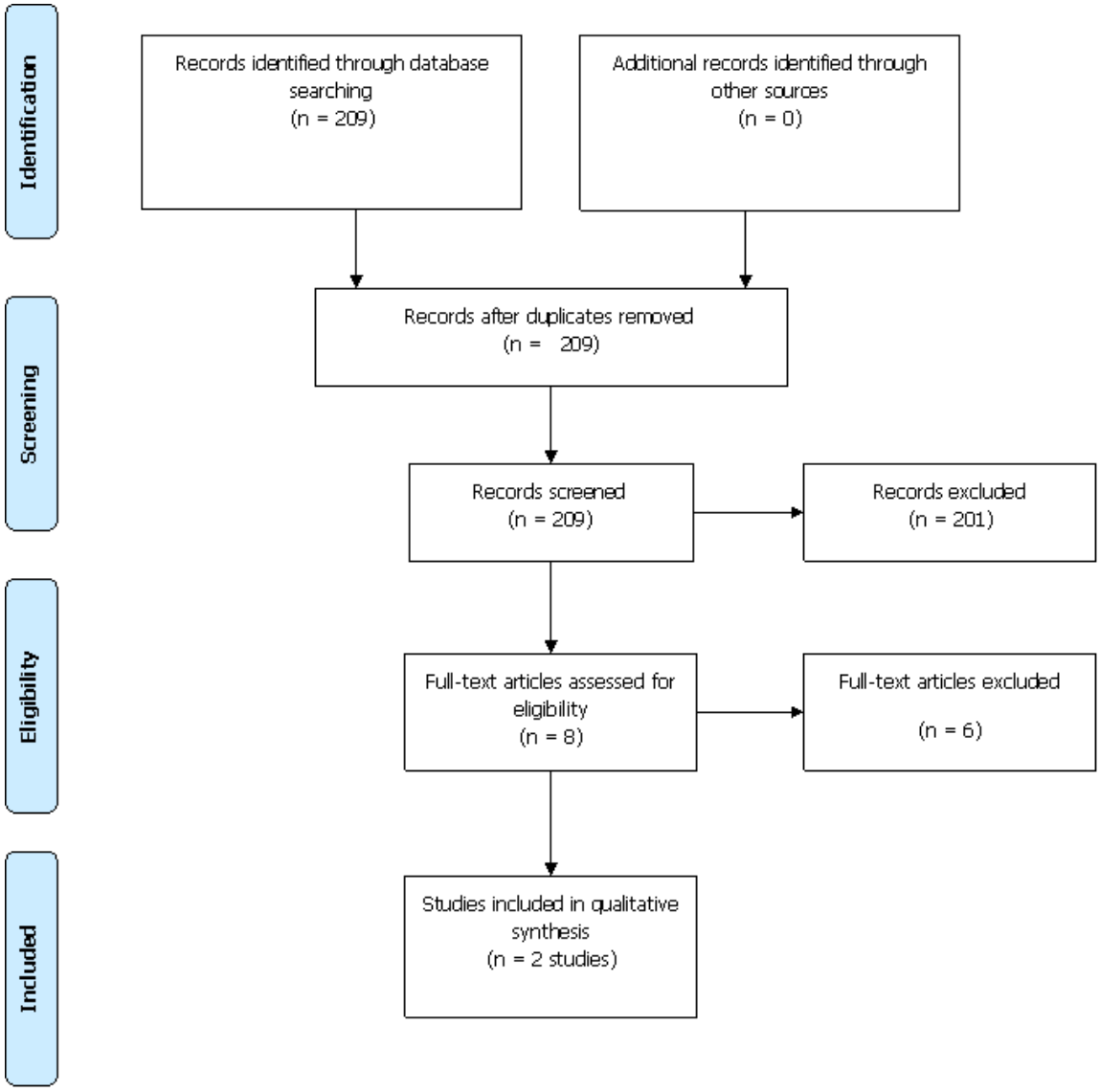
9.1 Trial Flow

The search of the literature yielded 209 separate references.

Following the first pass of the citations in the Heron SRDB, eight potentially relevant references were identified. Full-text reports of these citations were obtained for more detailed evaluation. Following detailed examination of the reports, six studies were excluded.

Two references reporting trials met the inclusion/exclusion criteria for this review. Data were extracted from two reports, and extraction grids were prepared for the two trials.

Figure 36: Flow diagram (economic review)



9.2 Overview of the identified studies

Two studies were identified which met the inclusion criteria for the review and are summarised in Table 62.

Remak 2008 was a Markov Model based study assessing the cost effectiveness and cost utility of sunitinib as a first-line treatment in metastatic RCC (mRCC) compared with interferon (IFN) and interleukin-2 (IL-2) from a US societal perspective. The model followed a hypothetical cohort of 1,000 patients with mRCC and documented clear-cell histology, radiographically measurable lesions, adequate organ functions and ECOG performance status of 0 or 1 over the patient's lifetime (10 years).

Mickisch 2009 was a decision analytical model based study evaluating the costs of managing AEs of bevacizumab in combination with IFN compared to sunitinib in the first-line treatment of mRCC in United Kingdom, Germany, Italy and France.

Table 62: Summary of the economic evaluations identified in the economic systematic review

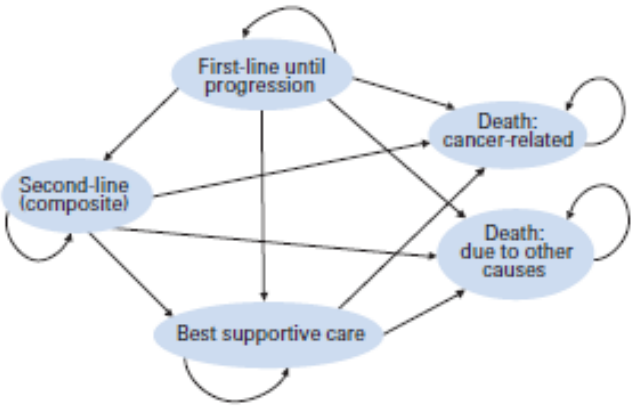
Study	Year	Country	Summary of model	Patient population (average year)	QALY (intervention comparator)	Costs (currency) (intervention, comparator)	ICER (per QALY gained)
(Remak 2008)	2006	US	The Markov model was developed in excel to simulate disease progression and outcomes over the lifetime (10 years) of a hypothetical cohort of 1000 patients with mRCC receiving first-line treatment with (in 6-week cycles) with sunitinib compared with IFN or IL-2.	Hypothetical cohort of 1000 patents	The estimated gains over IFN were 0.14 QALYs, and over IL-2 were 0.20 QALYs with sunitinib.	**Sunitinib** Drug cost (full dose): $5985/cycle Drug cost (reduced dose): $4488.75/cycle Cost of serious AEs: $160.13 **IFN** Drug cost (First cycle): $1903.10/cycle Drug cost (subsequent cycles): $2254.20/cycle Cost of serious AEs: $72.48 **IL-2** Drug costs: $13 903.54 Cost of serious AEs: $312.62	Incremental cost per PFY gained for sunitinib versus IFN was $18 611, and the ICER and ICUR of Sunitinib versus IFN were $67 215 per LY gained and $52 593 per QALY gained.
(Mickisch 2009)	Unclear	United Kingdom, Germany, Italy, France	An Excel-based linear decision analytical model was developed to calculate and compare the costs of management of all grades of AEs according to standard clinical practice for Bevacizumab plus IFN and sunitinib used as first-line treatment of metastatic RCC.	Not reported	Not reported	All grade AE management costs per patient for bevacizumab plus IFN arm in UK, Germany, and France were €1309, €1477, and €1957, respectively. All grade AE management costs per patient for Sunitinib arm in UK, Germany, and France were €2350, €2071, and €5127, respectively. Grade 3 – 4 AE management costs per patient for bevacizumab plus IFN and Sunitinib arm in Italy were €402 and €891, respectively.	Not reported

ICER = Incremental cost-effectiveness ratio, IL-2 = Interleukin-2, QALY = Quality adjusted life years.

9.3 Model design and inputs

Mickisch 2009 reported the results and inputs of an Excel-based linear decision analytic model for which the model structure was not presented. Remak 2008 however reported the results and inputs of a 10-year Markov model, for which the model structure is shown in Figure 37. In this model, patients were assumed to receive active treatment until an investigator's assessment of tumour progression was confirmed, then the patients were switched to either second-line treatment or to BSC.

Figure 37: Structure of the model presented in Remak 2008



Inputs into the economic models are presented below in Table 63. Both studies included direct medical costs including the treatment of AEs. Since the Mickisch 2009 study aim was to compare only the costs of managing AEs, efficacy and quality of life were not input into the evaluation, while Remak 2008 included both efficacy (PFS) and quality of life (EQ-5D).

Table 63: Overview of inputs into the economic evaluations

Input	Remak 2008	Mickisch 2009
Efficacy	Data from the second interim analysis of the pivotal phase III trial for sunitinib. Data from a randomised, multicentre, phase III study for IL-2.	-
QoL	EuroQol (EQ-5D) instrument was used for QoL data. Outcomes were valued in QALYs in accordance with economic assessment guidelines. Utility values from a phase II trial of second-line Sunitinib in mRCC were used to calculate utilities during second-line treatment and palliative care.	-
Safety	The model incorporated the following treatment-related AEs; fatigue/asthenia stomatitis, hypertension, thrombocytopenia, neutropenia, abnormal ejection fraction, nausea/vomiting, diarrhoea, anaemia, hand-foot syndrome, and infection. Resource use based on expert opinion and published sources.	The model used the total incidence of grade 1-4 AEs reported in phase III trials in the disease setting.
Costs	Direct medical costs included were; managing treatment-related serious AEs, diagnosis and treatment of progression, and BSC in the terminally ill. Indirect costs were not included. National average length of stay associated with each AE was based on the Agency for Healthcare	Costs included the management of AEs only. UK; from a review of published literature. Germany; calculated from the diagnosis-related group (DRG) funding system catalogue (2008) and from Einheitlicher Bewertungsmaβstab catalogue (2008). Costs include medicines and

	Research and Quality (AHRQ) Healthcare Costs and Utilization Project (HCUP) Nationwide inpatient Sample (NIS) database according to ICD-9 codes.	staff and maintenance. France; drug costs - Banque Claude Bernard database and from Pharmacie central des Hopitaux de Paris. Laboratory tests and examinations from official tariff lists. Hospitalisation costs estimated using French DRG hospital database and Etude Nationale de Couts. Italy; report of a Delphi panel of exports from five clinical practices, from Italian national DRG tariff and from two studies.

AE = Adverse Event, IL-2 = Interleukin-2, mRCC = metastatic renal cell carcinoma, QALY = Quality adjusted life years

9.4 Results of economic evaluations

The results of the economic evaluation presented by Remak et al (Remak 2008) are shown in Table 64. The results showed that sunitinib was both less costly and more effective than IL-2. In addition, sunitinib was more costly, but more effective than IFN, resulting in an ICER (LYs gained) of $67 215 and an ICUR of $52 593.

Table 64: Incremental cost-effectiveness and cost utility ratios for sunitinib versus IFN and IL-2 in the model presented in Remak 2008

Model Outcome	Deterministic mean per treatment strategy		
	Sunitinib	IFN	IL-2
Cost, $	224 970	217 436	228 411
Progression free years	0.92	0.1	0.57
Life-years	2.09	1.98	1.85
QALYs	1.33	1.19	1.13
ICER – progression free years gained, $		18 611	Dominated
ICER – LYs gained, $		67 215	Dominated
ICUR, $		52 593	Dominated

ICER = Incremental cost effectiveness ratio , ICUR =Incremental cost utility ratio, IFN =Interferon alpha, IL-2 = Interleukin-2

Mickisch 2009 reported that the average cost per patient of managing all-grade and grade 3-4 AEs varied across the countries assessed in the evaluation, and that the costs were higher for sunitinib than for Bevacizumab plus IFN, Table 65. The main cost drivers were lymphopenia, neutropenia, thrombocytopenia, leucopenia and fatigue/asthenia for sunitinib; and proteinuria, fatigue/asthenia, bleeding, anaemia and gastrointestinal perforation for Bevacizumab plus IFN.

Table 65: All-grade AE management costs per patient (Euros) in each of the countries (Mickisch 2009)

Country	Sunitinib	Bevacizumab plus IFN	Cost saving*
UK	€2 350	€1 309	€1 041 (44%)
Germany	€2 071	€1 477	€594 (29%)
France	€5 127	€1 957	€3 170 (62%)

*cost saving of Bevacizumab plus IFN compared to sunitinib. IFN = Interferon alpha

9.5 Critical appraisal of the identified studies

In the study Mickisch 2009, Costs and consequences were measured accurately and in appropriate physical units and the study examined both costs and effects of the treatments. The incremental cost analysis was not done in the study. All possible alternatives were explored through sensitivity analysis in this study. Remak et al.,

performed cost-effectiveness and cost-utility analysis and made conclusions on the basis of ICER and ICUR. The Markov model used in the study was well defined and sensitivity analyses of model parameters were performed. Both studies were critically appraised using the Drummond and Phillip checklist (see Appendix E).

10 Critique of models submitted to HTA agencies

10.1 Overview of HTA submissions identified

Several submissions were identified for the agencies AWMSG, CADTH, PBAC, NICE and the SMC in RCC, Table 66. In addition, a few submissions were for treatments intended for second-line therapy in patients who had failed previous cytokine therapy (or in those patients for which this treatment was not suitable); these studies are not discussed in the following sections.

Table 66: Summary of the HTA submissions identified in RCC

HTA agency	Intervention	Patient population	Date of guidance issued	Decision
AWMSG	Sorafenib	Advanced RCC in patients who have failed prior IFN-α or interleukin-2 based therapy, or are considered unsuitable for such therapy.	5 June 2007	Not recommended
	Sunitinib	Advanced and/or mRCC in first-line therapy as alternatives to cytokines.	15 August 2007	Not recommended
CADTH	Sorafenib	Advanced/metastatic RCC (clear cell) in patients who have failed prior cytokine therapy or are considered unsuitable for such therapy.	28 February 2007	Not recommended
	Sunitinib	Advanced RCC in patients who have failed prior IFN-α or interleukin-2 based therapy, or are considered unsuitable for such therapy.	26 April 2007	Not recommended
PBAC	Sorafenib	Initial (up to 3 months) treatment of advanced (unresectable or metastatic) RCC in patients with WHO performance status of 2 or less. Continuing treatment of advanced RCC (beyond 3 months) in patients with stable disease or responding disease (according to RECIST criteria).	November 2006	Not recommended
			March 2008	Not recommended
	Sunitinib	Advanced/metastatic RCC in patients with an ECOG performance status of 0 or 1.	March 2007	Decision deferred
			March 2008	Not recommended
		Stage IV RCC (clear cell) in patients with an ECOG performance status of 0 or 1.	July 2008	Recommended
	Temsirolimus	Advanced RCC in patients with a poor prognosis	July 2008	Not recommended
NICE	Bevacizumab	First-line therapy treatment options for people with advanced and/or metastatic RCC.	August 2009	Not recommended
	Sorafenib	Advanced and/or metastatic RCC in patients who have failed IFNIFN or IL-2 therapy or who are considered unsuitable for such therapy.	August 2009	Not recommended
	Sunitnib	First-line therapy treatment options for people with advanced and/or metastatic RCC.	March 2009	Recommended
		Second-line treatment options for people with advanced and/or metastatic RCC.	August 2009	Not recommended
	Temsirolimus	First-line therapy treatment options for people with advanced and/or metastatic RCC with at least 3 of 6 prognostic risk factors.	August 2009	Not recommended
SMC	Sorafenib	Advanced RCC in patients who have failed prior IFN or interleukin-2 based therapy or are considered unsuitable for such therapy.	6 October 2006	Not recommended
	Sunitinib	Advanced and/or metastatic RCC after failure of IFN or interleukin-2 therapy.	12 January 2007	Not recommended
	Sunitinib	Advanced and/or metastatic RCC.	8 June 2007	Not recommended
	Bevacizumab plus IFNBevacizumab plus IFN	First-line treatment of patients with advanced and/or metastatic RCC.	11 February 2008	Not recommended

AWMSG = All Wales Medicines Strategy Group, CADTH = The Canadian Agency for Drugs and Technology in Health, PBAC = Pharmaceutical Benefits Advisory Committee , NICE = National Institute for Health and Clinical Excellence, SMC = Scottish Medicines Consortium, HTA = Health Technology Assessment, RCC = Renal cell carcinoma

10.2 All Wales Medicines Strategy Group

One submission for sunitinib in the treatment of advanced/metastatic RCC to the AWMSG was identified. The submission was for first-line therapy as an alternative to cytokines. The manufacturers submitted a decision analytic model to determine the lifetime expected costs and benefits of treatment with sunitinib versus IFN. The model included three health states namely progression-free, progressed and dead.

The AWMSG did not recommend the use of sunitinib in this patient population since the estimated ICER was based on an interim analysis of data from one clinical trial, which the agency did not feel demonstrated the cost-effectiveness of the treatment. Table 67 details the criticisms of the economic evaluation in this submission.

Table 67: Overview of critique for the sunitinib submission to the AWMSG

	Details of the aspect of the model criticised	Criticisms made by the AWMSG
Choice of comparator(s)	• The AWMSG did not criticise the choice of comparator (IFN)	
Choice of analysis type	• The AWMSG did not criticise the analysis type used (cost utility).	
Data inputs	• The model used the data from an interim analysis of a single, ongoing phase III trial.	• The validity of the results is therefore dependent on factors including the design of the original trial and the number of withdrawals from the study. • The generalisability of the results to other settings and patient populations is therefore also a concern. • The utility values taken from this trial for use in the model were not compared with utilities from alternative sources.
	• The study population comprised a hypothetical cohort of patients diagnosed with advanced and/or mRCC with an ECOG performance status of 0 or 1 (trial A6181034).	• The outcomes in patients with a poorer performance status is uncertain. • It was not stated what percentage of the patients were UK-based.
	• The time horizon selected for the base case analysis was six years, based on results of extrapolating data from the first 15 months of trial A6181034.	• Considerable uncertainty was involved in extrapolating overall survival curves and this was acknowledged.
	• The model considered treatment-related adverse effects which were statistically significantly different between regimes, and, that were determined by expert opinion as likely to incur resource use costs.	• In the model supplied, the percentages of patients experiencing neutropenia, and vomiting, which fit both of these requirements, were not clearly reported. • The impact of uncertainty in the frequency of adverse effects was not investigated in the sensitivity analysis.
	• The drug acquisition costs used in the model were those obtained from the "new" costs published for 1st April 2007 and assumed there was no charge to the NHS for the first cycle of sunitinib.	• The effect of a change in the cost of sunitinib or IFN was not explored in the one-way or probabilistic sensitivity analysis even though it is highly likely to affect the results.
	• On the basis of clinical opinion, it was assumed that 50% of patients would require the assistance of a district nurse to administer their subcutaneous IFN dose, this was used to calculate the drug administration costs for IFN in the model.	• This figure appears high and was not varied in the sensitivity analysis. The higher rate of administration costs applied to IFN could potentially bias the results in favour of sunitinib.
Model design	• The model did not account for possible spontaneous remission.	• This is known to occur occasionally in untreated patients, but is not incorporated into the model.

	• Probabilistic sensitivity analysis (PSA) was conducted by assigning distributions to parameters of the base-case model. The PSA did not appear to assign distributions to the cost of both interventions.	• This did not seem appropriate to the AWMSG.
Model presentation	• The company submission did not include a diagram of the decision tree nor the disease progression model.	

10.3 Pharmaceutical Benefits Advisory Committee

10.3.1 Sorafenib

Sorafenib was submitted to PBAC on two occasions, first in November 2006 and the resubmission owing to an initial rejection from the agency, in March 2008. Sorafenib was submitted for use as initial therapy in patients with advanced RCC with a WHO performance status of 2 or less, and for continuing therapy in patients with stable disease or responding disease according to the RECIST criteria. Little detail regarding the economic evaluation was provided in the public summary document, however, some key criticisms were reported (Table 68).

Table 68: Overview of critique for the sorafenib submission to PBAC

	Sorafenib vs. Best supportive care	
	Details of the aspect of the model criticised	**Criticisms made by the PBAC**
Data inputs	• The key concern raised by the modelled economic evaluation related to the time horizon. The assumptions resulted in the model predicting an incremental survival in excess of the survival shown in the results of the clinical trial whose data was used for the model.	
Results	• The PBAC considered that the conservative base case incremental cost-effective ratio (ICER) estimate in the range $45,000 - $75,000 to be high and uncertain with the possibility of the ICER being greater than $150,000.	

In March 2008, the manufacturers of sorafenib resubmitted for use in the same patient population; "sorafenib is expected to be used alongside the current practice of best supportive care and in the minority of cases, it may replace immunotherapy or chemotherapy in the treatment of advanced RCC".

In the resubmission, a new modelled economic evaluation was presented, which differed from the original model in the following aspects;

- Duration 5 years instead of 11 years
- Modelled two health states (alive and dead) instead of three (progression-free, disease progression and dead)
- Used a more patient relevant outcome, survival, rather than progression.

Despite the changes to the model, sorafenib remained not recommended for use in this patient population, following further critique from PBAC, Table 69.

Table 69: Overview of critique for the sorafenib resubmission to PBAC

	Sorafenib vs. Best supportive care	
	Details of the aspect of the model criticised	**Criticisms made by the PBAC**
Data inputs	• In the previous submission (November 2006) the key concern raised by the modelled economic evaluation related to the time horizon where nine to twelve months' worth of data were extrapolated to eleven years. The model structure in the current re-submission had improved, with a shorter time horizon of 5 years.	• These changes were not enough to offset the clinical uncertainties.
	• Modelled two health states (alive and dead) instead of three	

	• The model used a more patient relevant outcome, survival, and thus avoided the problem of limited evidence on the patient relevance of "progression".	• The model relied on an outcome, overall survival, the result of which failed to attain statistical significance in the trial. • Issues of uncertainty remained from the extrapolation of the overall survival data from the clinical evidence

PBAC = pharmaceutical benefits advisory committee

10.3.2 Sunitinib

Sunitinib was submitted to PBAC in 2007 for use in the treatment of advanced/metastatic RCC in patients with an ECOG performance status of 0 or 1. However at the March 2007 meeting, the decision was deferred pending the provision of further economic analyses to demonstrate whether the treatment was suitably cost effective. The public summary document for this submission was not available.

At the March 2008 meeting, sunitinib was reconsidered for use in this patient population. In this submission a cost-effectiveness approach was presented. The Public Summary Document did not report detailed information regarding the manufacturer's evaluation or detailed criticism of this model.

The PBAC did not explicitly criticise any aspect of the economic evaluation submitted with this HTA submission. However, a few comments were made about the economic evaluation submitted:

- "The PBAC considered the submission's Markov model, which extrapolated overall survival from progression-free survival, to be more informative than the direct extrapolation of overall survival data requested during evaluation, in this instance.
- The direct extrapolation of overall survival, presented in the Pre-PBAC response, showed a two year survival gain which is implausible. The uncertainty of extrapolating overall survival with so few observations in this context outweighed the uncertainty associated with extrapolating survival from pooled mortality rates for progressors and non-progressors using progression-free survival rates.
- "The model inappropriately did not include costs or utilities for these AEs", such as heart failure and ischemia.

The PBAC considered the incremental cost effectiveness ratio of between $75,000 and $105,000 per extra QALY gained to be unacceptably high and uncertain, therefore did not recommend the use of sunitinib in this patient population.

In July 2008, sunitinib was considered for a third time by PBAC. The patient population for which the submission addressed was however reworded owing to previous concerns by PBAC; "treatment should be limited to clear cell disease" and "advanced should be replaced by stage IV disease".

The economic evaluation in this resubmission was unchanged, however addressed areas of uncertainty surrounding the ICER. Additional analyses were conducted to address these uncertainties. The following changes to the model were reported;

- Price was reduced (to be achieved via a risk sharing agreement)
- Pre and post-progression mortality rates derived from a Landmark analysis which was conducted to avoid guarantee-time bias in the survival estimates.

The revised estimated ICER was in the range of $45,000 – 75,000 per QALY and was "considered high but robust and acceptable in an area of high clinical need where no

effective alternative treatments are currently available". PBAC therefore recommended sunitinib for use in this patient population.

10.3.3 Temsirolimus

Temsirolimus was submitted to PBAC for use in advanced RCC in patients with a poor prognosis. The manufacturers presented a stepped economic evaluation, estimating the cost-effectiveness of the intervention compared to BSC and a comparison of the interventions with IFN in a stepped economic evaluation. The model was a Markov-like structure with three health states (no progression, progression and death), and a time horizon of 3 years (36 one-month cycles), and included only AEs of grade 3 or 4 severity. Key criticisms of the evaluation by PBAC were reported, which contributed to the rejection of the treatment in this patient population, Table 70.

Table 70: Overview of critique for the temsirolimus submission to PBAC

	Temsirolimus vs. Best supportive care/IFN	
	Details of the aspect of the model criticised	**Criticisms made by the PBAC**
Data inputs	• Lack of exchangeability across trials in indirect comparison to estimate hazard ratios – particularly differences in IFN (common arm) dosing and extended period of time (10 years) between when trials were undertaken. • Challenges in estimating mean survival benefit.	• The PBAC considered these to be main areas of uncertainty in the economic analysis.
	• Not reporting major sources of uncertainty within the trial period or their impact on cost effectiveness – e.g. not reporting a confidence interval for the temsirolimus-placebo treatment effect or absolute life years difference.	• The Pre-Sub-Committee Response stated that it was not possible to determine confidence intervals for the temsirolimus-placebo effect in terms of overall survival as the confidence interval associated with treatment effect of IFN over BSC were not reported.

PBAC = pharmaceutical benefits advisory committee

10.4 National Institute for Health and Clinical Excellence

A Multiple Technology Appraisal (MTA) was originally set up by NICE to evaluate bevacizumab, sorafenib, sunitinib and temsirolimus for the treatment of advanced and/or metastatic RCC. An STA was then created for sunitinib in first-line treatment in this patient population following the initial MTA submissions.

10.4.1 Bevacizumab

The manufacturer of bevacizumab submitted a state-transition model with three health states (progression free survival, progressive disease and death), and compared the intervention in combination with IFN to IFN alone. Bevacizumab was not recommended for use by NICE and key criticisms of the model are reported in Table 71.

Table 71: Overview of critique for the bevacizumab submission to NICE

	Bevacizumab + IFN vs. IFN	
	Details of the aspect of the model criticised	**Criticisms made by the Assessment Group**
Data inputs	• In the model, dose intensity data for bevacizumab was estimated using the average time taking the drug in the trial divided by the average time patients spend in PFS in the model. Although these data were not reported in the written submission they were used in the model.	• These data, applied to adjust costs, were different to that reported in the RCT upon which the model was based, and different to the data quoted in the Roche submission. • When the dose intensity data reported in the published RCT were used in the manufacturer submitted model, the base case ICER increased substantially (£75 000 per QALY to £117 000 per QALY).
	• The hazard ratio values for overall survival and progression free survival used in the clinical effectiveness part of the model came from unpublished data on what was classed a "safety population", not from the clinical effectiveness data of the RCT upon which the model was supposed to be based.	• It was not clear why the manufacturer analysis used data from the safety population (compared to RCT data). • When the hazard ratios for overall survival from the RCT data were used in the manufacturer's model instead of the data from the unpublished "safety population" the ICER increased (from £75,000 per QALY to £87,400 per QALY).
	• The manufacturer performed a (probabilistic sensitivity analysis) PSA, but not univariate sensitivity analysis on parameters.	• This was noted as a concern for the Assessment Group.
Model design	• The manufacturer used a single type of model.	• This was noted as a concern for the Assessment Group.
	• The manufacturer provided no evidence to indicate that the mathematical logic of the model had been tested.	• This was noted as a concern for the Assessment Group.
	• The results of the model were not calibrated against independent data, although it is not clear that such independent data exist. • The results of the model had not been compared with those of other models of metastatic RCC, although these other models were reported only in abstract form.	• This was noted as a concern for the Assessment Group.
Sensitivity analysis	• In sensitivity analysis the submission reported findings where cost-effectiveness had been assessed using a log-logistic model (instead of the Gompertz methods in the base case analysis).	• The appropriateness and prominence of this sensitivity analysis was questioned, the reviewers of the model did not see the log-logistic method as a credible approach (Roche acknowledged that the log-logistic model results in an expected lifetime may be unrealistically long).

ICER = Incremental cost-effectiveness ratio, QALY = Quality adjusted life years

10.4.2 Sunitinib

Sunitinib was submitted to NICE for its use in both the first-line and second-line treatment of advanced/metastatic RCC. Sunitinib was compared to IFN in a simple state transition model with three health states (progression free survival, progressed disease and death). This model was critiqued, an overview of which is presented in Table 72.

Table 72: Overview of critique for the sunitinib submission to NICE

	Details of the aspect of the model criticised	Criticisms made by the Assessment Group
Data inputs	• The Weibull curve fitted to trial data on progression free survival used in the model for IFN is a poor fit to the empirical survival data.	• The Assessment Group considered the consequences of this poor fit to be important in that the modelling creates an underestimate of the benefits of IFN and an underestimate of the cost per QALY for sunitinib.
	• The manufacturer model assumes that people receive sunitinib or IFN until disease progression.	• Based on the views of the expert advisory group, the Assessment Group felt that IFN will generally be prescribed for a maximum period of 12 months. Therefore, the model may overestimate the costs and effects associated with IFN treatment (i.e. underestimate the incremental cost for sunitinib).
	• The data used for the model estimates that 50% of patients self-inject, and that the remainder have injections given by a district nurse at home, at a cost of £21 per visit.	• The Assessment Group felt that a higher proportion may self administer; therefore the submission probably slightly overestimates the cost of IFN.
	• Health state utilities/values data used in the model are unpublished.	• This was noted as a concern for the Assessment Group.
	• There appeared to be some potentially mixed numbers for the standard errors of hazard ratio figures used in the model (standard errors of the hazard ratios for overall survival are used for progression free survival and vice versa.	• Confusion in the assignment of data will affect the results of the probabilistic sensitivity analysis (PSA).
Model design	• The manufacturer used a single type of model.	• This was noted as a concern for the Assessment Group.
	• No evidence has been presented to indicate that the mathematical logic of the model has been tested.	• This was noted as a concern for the Assessment Group.
Sensitivity analysis	• The base case analysis demonstrated an ICER using the full ITT population of £82,003 per QALY. • The sensitivity analysis performed by the manufacturer for the economic model used separate source for the progression free survival and overall survival data to predict baseline IFN progression.	• The Assessment Group felt that using different data sources for OS and PFS in the model has the consequence/potential to distort the modelled disease progression due to the fact that the number of people in the progressive disease health state over time is calculated from (is a function of) related data on PFS and OS.
End of life criteria		• The committee was aware that the total number of people with advanced and/or metastatic RCC in England and Wales was approximately 4000. It therefore considered that for this appraisal, sunitinib should be regarded as meeting this criterion for an end-of-life treatment. • The Committee noted the normal life expectancy with IFN-α treatment ralrely exceeded 24 months and was potentially as low as 12 months. The sunitinib trial demonstrated sunitinib increased survival by more than 3 months in comparison with IFN-α alone. • It was further persuaded that sunitinib provided a step-change in the first-line treatment of advanced and/or metastatic RCC and noted that more than 20% of the public and patients that responded in consultation highlighted this impressive benefit from sunitinib. In summary, • The Committee was satisfied that sunitinib currently meets the criteria for being a life-extending end-of-life treatment, and that the evidence presented for this consideration was sufficiently robust.

IFN = Interferon alpha, OS = Overall survival, PSA =Probabilistic sensitivity analysis

10.4.3 Temsirolimus

A state-transition model with three health states (progression free survival, post-progression and death) was submitted for temsirolimus. The model compared the

intervention to IFN. Key criticisms of the model are presented in Table 73.

Table 73: Overview of critique for the temsirolimus submission to NICE

	Details of the aspect of the model criticised	Criticisms made by the CDR
Data inputs	• The manufacturer model assumes all IFN is administered in the hospital outpatient setting, costing £127.80 per visit.	• Based on information on current practice from the expert advisory group, the Assessment Group did believe this is an accurate reflection of current practice, who expected that in most cases IFN injections would be administered in the patient's home either by themselves or by friends, relatives or carers. If administration data from the expert advisory group is used in the manufacturer's model, the ICER increases substantially (from £55 814 per QALY to £102 000 per QALY).
	• The cost data used in the model assumes that the drug administration costs for temsirolimus should be adjusted using dose intensity data from the key RCT (costs are reduced). However, this assumption is not applied to costs associated with IFN.	• This was noted as a concern for the Assessment Group, when the suggestions on drug cost provided by the expert advisory panel were used in the manufacturer's model, the ICER of temsirolimus versus IFN increases substantially (from £55 800 to £74 819 per QALY).
	• The data used to derive health state utilities are not published.	• This lack of transparency was noted as a concern for the Assessment Group.
Model design	• The manufacturer used a single type of model.	• This was noted as a concern for the Assessment Group.
	• The results of the model were not calibrated against independent data.	• The model predictions of progression free survival and overall survival curves do not agree with the Kaplan-Meier curves reported in the key RCT. Given the expectation that the cost-effectiveness estimates are sensitive to the shape of the PFS and OS survival curves, this is an important discrepancy. • The results of the model have not been compared with those of other models of metastatic RCC, although these other models have been reported only in abstract form.
	• No evidence has been presented to indicate that the mathematical logic of the model has been tested.	• This was noted as a concern for the Assessment Group.

10.5 Scottish Medicines Consortium

10.5.1 Sunitinib

The SMC did not recommend the use of sunitinib for the treatment of advanced and/or metastatic RCC, following a submission to the agency in 2007. The manufacturer submitted a cost-utility analysis with three health states; alive and progression-free, alive and progressed, and dead. Patients could not switch between treatments when disease progressed and instead were offered best supportive care. The duration of the model was six years, and given the short term nature of the available trial data, extrapolation was required to estimate survival over this time period.

The notice of advice stated that the "manufacturer did not present a sufficiently robust economic analysis", and the key criticisms of the economic evaluation are presented in Table 74.

Table 74: Overview of critique for the sunitinib submission to SMC

	Details of the aspect of the model criticised	Criticisms made by the SMC
Choice of comparator(s)	• The SMC did not criticise the choice of comparator (IFN).	
Choice of analysis type	• The SMC did not criticise the analysis type used (cost utility).	
Data inputs	• The model used outcomes data from short term data extrapolated over 6 years.	• For the overall survival outcome, the data are sparse at this stage. Therefore there is uncertainty in knowing the true magnitude of the survival advantage of sunitinib over IFN alpha in the extrapolated phase of the model • Statistical advice received by the SMC on the progression-free and overall survival estimates suggested that they originated from a model with poorly fitted data. • The results were sensitive to changes in the values of these parameters. A probabilistic analysis indicated a 36% chance of sunitinib being cost effective at a willingness to pay for a QALY of £30 000.

SMC = Scottish Medicines Consortium

10.5.2 Bevacizumab

Bevacizumab was submitted to the SMC for use in combination with IFN for the first-line treatment of patients with advanced and/or metastatic RCC. The relevant Statement of Advice noted that the holder for the marketing authorisation for bevacizumab did not make a submission to the SMC regarding this product in this indication. Therefore the intervention is not recommended for use in NHS Scotland.

11 Discussion

11.1.1 Clinical Review

The objective of this review was to determine the efficacy and safety of pazopanib in comparison to sunitinib, sorafenib, bevacizumab plus IFN, temsirolimus, IFN and IL-2 for advanced/metastatic renal cell carcinoma in treatment naïve patients.

A total of 13 trials reported in 86 publications were extracted and the data was analysed. Two trials (Soret 1996, Prummer 1994) published in four publications met other inclusion/exclusion criteria but did not report data of interest to the review and were thus excluded from qualitative and quantitative analysis.

One RCT examining the clinical efficacy and safety of pazopanib versus placebo/BSC (VEG105192) was identified and included in the analyses

A second pazopanib study was identified during the review progress but this was excluded from the final list of included studies owing to the fact that (i) it was designed as a randomised discontinuation study but was later revised to a single-arm open-label study and (ii) lack of outcome data for treatment-naïve patients in the randomised phase of the study (VEG102616). Results from this study are presented in this document for completeness but they did not contribute to any analyses.

Head-to-head comparisons of pazopanib versus the other interventions of interest were not available. Hence, indirect comparison served as the only option to determine the relative efficacy and safety of the included interventions.

Pazopanib, during direct comparison with placebo (VEG105192), demonstrated a statistically significant improvement in PFS and overall response rate in treatment-naïve patients with advanced/metastatic RCC. Regarding the interim OS analysis, overall survival also appeared to be prolonged in the pazopanib arm relative to the placebo arm.

Pazopanib demonstrated comparable PFS and OS to sunitinib, sorafenib and bevacizumab plus IFN

A solely treatment-naïve population was conscripted in all of the trials of interest except for two (VEG105192 and Target Study). These two studies included a mixed population of treatment-naive and pre-treated patients but reported subgroup data specifically for the treatment-naive sub-populations. The majority of the patients had ECOG/ WHO performance status 0 or 1. Numerically, similar median PFS was seen to sunitinib and superior PFS to sorafenib and bevacizumab plus IFN.

Indirect comparisons of PFS between pazopanib and the other interventions showed that pazopanib had comparable efficacy to sunitinib with no statistically significant difference. It showed favourable hazard ratios (lower risk) over sorafenib and bevacizumab plus IFN; however, the differences were not statistically significant.

Additionally, when pazopanib was indirectly compared with other active comparators (sunitinib or bevacizumab plus IFN), it exhibited similar efficacy with respect to overall survival without any statistically significant differences.

Similar overall response rates are observed for pazopanib, sunitinib, sorafenib and bevacizumab plus IFN

The results of meta-analysis showed that overall response rates were higher for pazopanib, sorafenib, IFN and IL-2 compared to placebo/BSC and all improvements over placebo were statistically significant, with the exception of IL-2. Pazopanib, when

compared indirectly with sunitinib or sorafenib or bevacizumab plus IFN, showed numerically lower response rates (irrespective of assessment by IRC or investigator); however, these differences were not statistically significant.

The relative efficacy of pazopanib and temsirolimus could not be determined by indirect comparison owing to significant differences in trial populations

Indirect comparisons of temsirolimus and pazopanib were possible for certain efficacy outcomes, though were not conducted, owing to the significantly different patient populations. Although there was an overlap in the patient populations in the two trials, more patients in the temsirolimus trial were MSKCC poor risk than those enrolled in the pazopanib trial. Since the patient populations were not comparable, and no sub-group data for the MSKCC intermediate risk group in the temsirolimus trial were available, an indirect comparison between the treatments was not conducted, as the heterogeneity between trials would not result in a true representation of the differences in efficacy between treatments.

As expected, due to the nature of the treatments, all treatments were associated with a high proportion of patients experiencing any AE

Pazopanib, sorafenib, bevacizumab plus IFN and temsirolimus were all associated a high proportion of patients, generally >90% experiencing an AE of any grade. Data for sunitinib for this outcome were not available from the identified study. The majority of these AEs were mild (grade 1 or 2) since generally less than 50% of patients experienced a grade 3 or 4 AE.

When indirectly comparing pazopanib with the other interventions of interest, pazopanib showed improved safety (lower risk of any AE/any grade 3 or 4 AE) over sorafenib and bevacizumab plus IFN, although these differences were not statistically significant. Due to a lack of sunitinib data for this outcome, an indirect comparison with pazopanib was not possible.

Pazopanib had a favourable safety profile versus sunitinib, sorafenib, temsirolimus and bevacizumab plus IFN, particularly in terms of blood and lymphatic disorders, including anaemia

After indirect comparisons, pazopanib showed a better safety profile over sorafenib and bevacizumab plus IFN, however, the difference was not statistically significant. Pazopanib was associated with a reduced risk of almost all AEs for which comparisons could be conducted compared to sunitinib. These included diarrhoea, vomiting, fatigue, hand-foot syndrome, total bilirubin increased, anorexia, epistaxis, and haematological AEs. The difference in risk, however, rarely reached statistical significance. Alopecia was the only AE where pazopanib showed an increased but statistically insignificant risk compared to sunitinib. On qualitative analysis, it was observed that higher proportion of patients had increased AST and ALT levels after the treatment with pazopanib as compared with sunitinib.

Across all the treatments examined in the indirect analysis, pazopanib demonstrated comparable, and for certain AEs, improved safety. This was particularly evident for the blood and lymphatic disorder AEs such as anaemia, where the risk of these events was consistently lower with pazopanib than the comparators, albeit these differences did not reach statistical significance.

The qualitative findings also suggested that pazopanib had an improved safety profile, particularly with regard to haematological side effects than sunitinib; the only exception was hair colour change. Pazopanib demonstrated relatively low proportions of patients experiencing any grade specific AEs, while sunitinib demonstrated much higher percentages of patients experiencing these events.

Within-group improvement from baseline in QoL outcomes demonstrated non-significant differences between pazopanib and placebo, however between group differences using endpoint values were not explored

A significant benefit of sunitinib over IFN was observed for QoL in the Cella 2008 study, while no significant difference was observed for pazopanib compared to placebo. However, the pazopanib study compared the mean difference from baseline in EQ-5D scores for pazopanib and placebo, while Cella 2008 compared only endpoint scores between treatments. The authors of the sunitinib study also report that "results predominately reflected between-group differences rather than within-group improvement from baseline", therefore may not have demonstrated significant results given the same analysis as that conducted in the pazopanib trial.

Furthermore, specific AEs were generally mild (<25% achieving grade 3 or 4 AEs), and for many specific AEs, no patients on pazopanib experienced a grade 3/4 AE, highlighting its improved safety profile versus comparators

As previously stated the majority of AEs experienced by patients in these studies were mild, with generally less than 50% of patients experiencing a grade 3 or 4 AE. The proportions of patients experiencing a specific grade 3 or 4 AE were generally low (1-25%) and comparable between sunitinib, bevacizumab plus IFN, temsirolimus and sorafenib. However, many cases were observed where none of the patients being treated with pazopanib experienced a grade 3 or 4 AE, including the events of mucositis, hand-foot syndrome and altered taste, suggesting an improved safety profile over the comparators.

Several data limitations were identified including the lack of head-to-head comparisons and the absence of data for certain AEs (e.g. hair colour change) in the IFN trials that were key to allowing indirect comparison between pazopanib and sunitinib

- One of the limitations with the data available was the lack of head-to-head comparisons between pazopanib and the other included interventions. Another factor limiting the analyses was the lack of data reported in some studies. For example, analyses of time to response and duration of response were limited by several of the studies not reporting these outcomes.
- Patients in the temsirolimus trial had a poorer prognosis than those in the other included studies, preventing indirect analysis of pazopanib with temsirolimus.
- There are several AEs (e.g. mucositis) which have become of increased clinical importance and focus since the introduction of the targeted agents. The IFN studies, which provide a critical path to the indirect comparison of these treatments to pazopanib, are older studies which do not report data for these AE outcomes. Therefore the relative risk of patients experiencing these events compared to pazopanib could not be established.

Pazopanib is an effective treatment for treatment-naïve patients with advanced/metastatic renal cell carcinoma, demonstrating increased efficacy compared to placebo and comparable efficacy to the current standard of care, sunitinib.

Pazopanib also demonstrated an improved safety profile, particularly for hematological AEs, over sunitinib.

11.1.2 Economic Review

Only two economic evaluations were identified in the review, one of which was a US based study and the other was conducted in multiple European countries, including the UK

One study was a Markov Model based study assessing the cost effectiveness and cost utility of sunitinib as a first-line treatment in metastatic RCC compared with IFN and IL-2 from a US societal perspective.

The second was a decision analytical model based study evaluating the costs of managing AEs of bevacizumab in combination with IFN compared to sunitinib in the first-line treatment of metastatic RCC in United Kingdom, Germany, Italy and France.

One model demonstrated that sunitinib was both less costly and more effective than IL-2, while the second study reported that the average cost per patient of managing all-grade and grade 3/4 AEs varied across the European countries, and that the costs were higher for sunitinib than for bevacizumab plus IFN

- The results of the economic evaluation presented by Remak 2008 showed that sunitinib was both less costly and more effective than IL-2.
- In addition, sunitinib was more costly, but more effective than IFN, resulting in an ICER (LYs gained) of $67 215 and an ICUR of $52 593.
- Mickisch 2009 reported that the average cost per patient of managing all-grade and grade 3/4 AE varied across the countries assessed in the evaluation, and that the costs were higher for sunitinib than for bevacizumab plus IFN.
- The main cost drivers were lymphopenia, neutropenia, thrombocytopenia, leucopenia and fatigue/asthenia for sunitinib; and proteinuria, fatigue/asthenia, bleeding, anaemia and gastrointestinal perforation for bevacizumab plus IFN.

Appendix A Study protocol

A.1 Clinical systematic review

Objectives and research questions	
Primary study question	• What is the relative clinical efficacy, safety and tolerability of pazopanib and other pharmacological interventions in the treatment of advanced/metastatic renal cell carcinoma in treatment-naïve patients?
Secondary study questions	• What is the effect of pazopanib and other pharmacological interventions on quality of life endpoints in advanced/metastatic renal cell carcinoma in treatment-naïve patients?
Studies to include	
Study designs	• Randomised control trials (RCTs) with any blinding status
Population	• Age: Adults (≥ 18 years) • Gender: Any • Race: Any • Stage of disease: Advanced and Metastatic (stage III/IV) • Treatment Naïve
Interventions	• Pazopanib monotherapy • IFN-α or IL-2 monotherapy • Bevacizumab (in combination with IFN-α) • Sunitinib monotherapy • Sorafenib • Temsirolimus
Comparator	• Any of the included interventions • Placebo • Best supportive care
Language	English only
Publication timeframe	1980 onwards for literature searches, last 3 years for hand searches
Other inclusion/exclusion criteria	-
Data sources	
Databases	• Medline • Embase • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Cochrane Methodology Register
Conference proceedings	• American Society of Clinical Oncology (ASCO) • ASCO-Genitourinary (ASCO-GU) • European Society for Medical Oncology (ESMO) • European Conference for Clinical Oncology (ECCO)
Other data sources	• Reference lists of previous trials and systematic reviews • Relevant websites for identification of ongoing trials (e.g. ClinicalTrial.gov; NCI clinical trial database; ISRCTN Register; UKCCCR Register of Cancer Trials; EORTC; UK Clinical Trials Gateway (UKCTG); metaRegister ((mRCT) of Controlled Trials)
Information to extract	
Study information	• Treatment arms • Treatment dose • Number of patients • Trial length • Method of randomisation • Study quality
Baseline data	• Age • Gender • Cancer stage of treatment arms • Prognosis • Performance status (ECOG or KPS) • MSKCC score • Prior nephrectomy • Prior treatment (in particular cytokine treatment) • Proportion of patients with each subtype of RCC • Proportion of patients who reduce the dose of treatment
Efficacy outcomes	• Overall survival • Progression free survival • Overall response rate (complete + partial responses) • Health-related quality of life

	• Time to progression • Proportion of patients with stable disease • Response duration • Time to response
Safety outcomes	• Adverse effects (AE) including incidence and severity (grade) of all AEs reported • Total withdrawals • Withdrawals due to AEs • Withdrawals due to death • Serious AEs
Other outcomes	
Analyses	
• Where possible, meta-analysis will be used to pool results for each outcome of interest and for each combination of treatment • Both fixed-effects (Mantel-Haenszel) and random-effects (DerSimonian and Laird) models will be used. • Continuous outcomes will be assessed as non-standardized mean differences with 95% confidence intervals • Dichotomous outcomes will be assessed as risk ratios with 95% confidence intervals	
Reporting	
This study is to be written up HTA-compliant systematic review so as to support STA submission.	

Best supportive care: no active treatment/observation/a method of care that is not a focused treatment/treatments which clinicians consider to be "placebo-equivalent" including medroxyprogresterone and vinblastine.

A.2 Economic systematic review

Objectives and research questions	
Primary study question	• What is the relative clinical efficacy, safety and tolerability of pazopanib and other pharmacological interventions in the treatment of advanced/metastatic renal cell carcinoma in treatment-naïve patients?
Secondary study questions	• What is the effect of pazopanib and other pharmacological interventions on quality of life endpoints in advanced/metastatic renal cell carcinoma in treatment-naïve patients?
Studies to include	
Study designs	• Randomised control trials (RCTs) with any blinding status
Population	• Age: Adults (≥ 18 years) • Gender: Any • Race: Any • Stage of disease: Advanced and Metastatic • Treatment Naïve
Interventions	• Pazopanib • Immunotherapy (IFN-α or IL-2) without the addition of bevacizumab • Bevacizumab (in combination with IFN-α) • Sunitinib • Sorafenib • Temsirolimus
Language	English only
Publication timeframe	1980 onwards for literature searches, last 3 years for hand searches
Other inclusion/exclusion criteria	
Data sources	
Databases	• MEDLINE • EMBASE • Cochrane Economic Evaluations Database • Cochrane Technology Assessments Database • Database of Abstracts of Reviews of Effects (DARE)
Other sources	• NICE • SMC • PBAC • CADTH • AWMSG
Information to extract	
Study information	• Treatment arms • Treatment dose • Number of patients • Trial length • Method of randomisation • Study quality
Patient population	• Age • Gender

	• Cancer stage of treatment arms • Prognosis • Prior treatment (in particular cytokine treatment) • Proportion of patients will clear cell carcinoma] • Down dosing
Model structure	• Model design • Model structure
Cost-effectiveness outcomes	• Cost-effectiveness of treatment
Other outcomes	• Perspective • Country
Analyses	
• A table of results of the models will be presented in the report • Models will be analysed according to the line of therapy • Any models in particular subgroups of patients will be highlighted and discussed separately	
Reporting	
This study is to be written up HTA-compliant systematic review so as to support STA submission.	

Best supportive care: no active treatment/observation/a method of care that is not a focused treatment/treatments which clinicians consider to be "placebo-equivalent" including medroxyprogresterone and vinblastine.

Appendix B Search strategy

B.1 Clinical systematic review

The search strategies for the clinical systematic review are presented in the following sections.

B.1.1 MEDLINE and Embase

Date search run: 23 November 2009

Table 75: Search for randomised controlled trials – Embase.com

#	Search History	Results
1.	'clinical trial'/exp	754474
2.	'randomization'/de	48076
3.	'controlled study'/de	3084530
4.	'comparative study'/de	573453
5.	'single blind procedure'/de	11501
6.	'double blind procedure'/de	92524
7.	'crossover procedure'/de	25892
8.	'placebo'/de	156832
9.	'clinical trial' OR 'clinical trials'	856289
10.	'controlled clinical trial' OR 'controlled clinical trials'	352414
11.	'randomised controlled trial' OR 'randomised controlled trial' OR 'randomised controlled trials' OR 'randomised controlled trials'	268489
12.	'randomisation' OR 'randomization'	59957
13.	Rct	5404
14.	'random allocation'	1000
15.	'randomly allocated'	12960
16.	'allocated randomly'	1589
17.	allocated NEAR/2 random	739
18.	(single OR double OR triple OR treble) NEAR/1 (blind* OR mask*)	152930
19.	placebo*	225692
20.	'prospective study'/de	135910
21.	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20	4174839
22.	'case study'/de	7268
23.	'case report'	1666932
24.	'abstract report'/de	89415
25.	'letter'/de	645194
26.	#22 OR #23 OR #24 OR #25	2270904
27.	#21 NOT #26	4066688
28.	'pazopanib'/de OR 'sunitinib'/de OR 'sorafenib'/de OR 'bevacizumab'/de OR 'temsirolimus'/de OR 'everolimus'/de OR 'interleukin 2'/de OR 'alpha interferon'/de	94250

#	Search History	Results
29.	'alpha-interferon':ab,ti OR alfaferone:ab,ti OR alferon:ab,ti OR 'alpha ferone':ab,ti OR cilferon:ab,ti OR ginterferon:ab,ti OR 'interferon-alpha':ab,ti OR introma:ab,ti OR kemron:ab,ti OR leukinferon:ab,ti OR leukinferron:ab,ti OR 'leukocyte interferon':ab,ti OR 'refecon a':ab,ti OR 'referon a3':ab,ti OR sumiferon:ab,ti OR sumipheron:ab,ti OR veldona:ab,ti	10766
30.	'biotest':ab,ti OR bioleukin:ab,ti OR 'interleukin-ii':ab,ti OR 'interleukin-2':ab,ti OR 'il-2':ab,ti OR il2:ab,ti OR 'ro-236019':ab,ti OR tcgf:ab,ti OR tsf:ab,ti	56840
31.	everolimus:ab,ti OR afinitor:ab,ti OR certican:ab,ti OR 'nvp-rad-001':ab,ti OR 'rad-001':ab,ti OR 'rad 001a':ab,ti OR rad001:ab,ti OR rad001a:ab,ti OR 'sdz rad':ab,ti	853
32.	temsirolimus:ab,ti OR 'cci-779':ab,ti OR 'cell-cycle-inhibitor-779':ab,ti OR 'nsc 683864':ab,ti OR nsc683864:ab,ti OR torisel:ab,ti	402
33.	bevacizumab:ab,ti OR avastin:ab,ti OR 'nsc 704865':ab,ti OR nsc704865:ab,ti OR 'anti-vegf':ab,ti OR 'rhumab-vegf':ab,ti	4100
34.	'bay 43-9006':ab,ti OR 'bay 439006':ab,ti OR 'bay43-9006':ab,ti OR bay439006:ab,ti OR nexavar:ab,ti OR sorafenib:ab,ti	996
35.	sunitinib:ab,ti OR sutent:ab,ti OR 'pha 2909040ad':ab,ti OR 'pha2909040ad':ab,ti OR 'su 010398':ab,ti OR 'su 011248':ab,ti OR 'su 10398':ab,ti OR su10398:ab,ti OR 'su 11248':ab,ti OR su010398:ab,ti OR 'su011248':ab,ti OR su11248:ab,ti	960
36.	armala:ab,ti OR pazopanib:ab,ti OR gw786034*:ab,ti OR (gw NEXT/1 786034*):ab,ti OR (sb NEXT/1 710468*):ab,ti OR sb710468*:ab,ti OR votrient:ab,ti	39
37.	#28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36	121860
38.	'kidney carcinoma'/de	27437
39.	'kidney tumour'/exp	64633
40.	renal*:ab,ti OR kidney*:ab,ti OR grawit*:ab,ti OR hypernephroid*:ab,ti OR nephroid*:ab,ti	602660
41.	carcinoma*:ab,ti OR cancer*:ab,ti OR neoplasm*:ab,ti OR adeno*:ab,ti OR tumo?r*:ab,ti OR pyelocarcinoma*:ab,ti OR metastas?s:ab,ti OR oncocytoma:ab,ti	1586263
42.	#40 AND #41	69298
43.	(metanephric NEAR/2 adeno*):ab,ti	136
44.	rcc:ab,ti OR mrcc:ab,ti OR 'm-rcc':ab,ti	5722
45.	'hypernephroma':ab,ti	1196
46.	#38 OR #39 OR #42 OR #43 OR #44 OR #45	100246
47.	#27 AND #37 AND #46	3814
48.	**#27 AND #37 AND #46 AND [1980-2010]/py**	**3884**

B.1.2 Cochrane

Date search run: 23 November 2009

Table 76: Cochrane library search strategy

ID	Search History	Results
#1	MeSH descriptor Interferon-alpha explode all trees	2099
#2	MeSH descriptor Interleukin-2 explode all trees	702
#3	("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):ab,ti,kw	3001
#4	(biotest OR bioleukin OR "interleukin-ii" OR "interleukin-2" OR "il-2" OR il2 OR "ro-236019" OR tcgf OR tsf):ab,ti,kw	1902

ID	Search History	Results
#5	(everolimus OR afinitor OR certican OR "nvp-rad-001" OR "rad-001" OR "rad 001a" OR rad001 OR rad001a OR "sdz rad"):ab,ti,kw	154
#6	(temsirolimus OR "cci-779" OR "cell-cycle-inhibitor-779" OR "nsc 683864" OR nsc683864 OR torisel):ab,ti,kw	25
#7	(bevacizumab OR avastin OR "nsc 704865" OR nsc704865 OR "anti-vegf" OR "rhumab-vegf"):ab,ti,kw	236
#8	("bay 43-9006" OR "bay 439006" OR "bay43-9006" OR bay439006 OR nexavar OR sorafenib):ab,ti,kw	63
#9	(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):ab,ti,kw	37
#10	(armala OR pazopanib OR gw786034* OR sb710468* OR votrient):ab,ti,kw	2
#11	(#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10)	5875
#12	MeSH descriptor Carcinoma, Renal Cell explode all trees	301
#13	(renal* OR kidney* OR grawit* OR hypernephroid* OR nephroid*):ab,ti,kw	24198
#14	(carcinoma* OR cancer* OR neoplasm* OR adeno* OR tumo?r* OR pyelocarcinoma* OR metastas?s OR oncocytoma):ab,ti,kw	60577
#15	(#13 AND #14)	1811
#16	(metanephric adj2 adeno*):ab,ti,kw	0
#17	(rcc OR mrcc OR "m-rcc"):ab,ti,kw	168
#18	hypernephroma:ab,ti,kw	4
#19	(#12 OR #15 OR #16 OR #17 OR #18)	1829
#20	(#11 AND #19)	334
#21	**(#11 AND #19), from 1980 to 2009 [Cochrane review, clinical trials, Method studies]**	**317**

B.1.3 MEDLINE in-process (2009 only)

Date search run: 2 December 2009

#15	Search ("2009/01/01"[Publication Date] : "3000"[Publication Date]) AND (#10 AND #13)	485
#14	Search #10 AND #13	5015
#13	Search #11 OR #12	200717
#12	Search ((((("Sutent"[Title/Abstract]) OR ("Votrient"[Title/Abstract])) OR ("Afinitor"[Title/Abstract])) OR ("Torisel"[Title/Abstract])) OR ("Nexavar"[Title/Abstract])	105
#11	Search (((((((("Pazopanib"[Title/Abstract]) OR ("Bevacizumab"[Title/Abstract])) OR ("Sunitinib"[Title/Abstract])) OR ("Temsirolimus"[Title/Abstract])) OR ("Interferon"[Title/Abstract])) OR ("interleukin"[Title/Abstract])) OR ("Everolimus"[Title/Abstract])) OR (Avastin)) OR (Sorafenib)	200712
#10	Search #8 OR #9	61128
#9	Search ((((("RCC"[Title/Abstract]) OR ("MRCC"[Title/Abstract])) OR ("M-RCC"[Title/Abstract])) OR ("hypernephroma"[Title/Abstract])) OR ("metanephric adenocarcinoma"[Title/Abstract])	6497
#8	Search #4 AND #7	59700
#7	Search #5 OR #6	1437992
#6	Search "oncocytoma"[Title/Abstract]	1286
#5	Search (((((((("carcinoma"[Title/Abstract]) OR ("cancer"[Title/Abstract])) OR ("neoplasm"[Title/Abstract])) OR ("adenocarcinoma"[Title/Abstract])) OR ("tumour"[Title/Abstract])) OR ("tumour"[Title/Abstract])) OR ("pyelocarcinoma"[Title/Abstract])) OR (metastasis)) OR (metastases)	1437614
#4	Search #1 OR #2 OR #3	512295
#3	Search "nephroid"[Title/Abstract]	11

#2	Search ("grawit"[Title/Abstract]) OR ("hypernephroid"[Title/Abstract])	210
#1	Search ("renal"[Title/Abstract]) OR ("kidney"[Title/Abstract])	512199

B.1.4 Meta-register search

Date search run: 2 December 2009

Search term: (Pazopanib OR Bevacizumab OR Sunitinib OR Temsirolimus OR Interferon OR interleukin OR Everolimus OR Sorafenib OR Avastin OR Sutent OR Nexavar OR Torisel OR Afinitor OR Votrient)

Limit: UKCTG, ISRCTN

Total retrieved: 153

B.1.5 Clinicaltrial.gov search

Date search run: 2 December 2009

Search term:

Search strategy	Search result
Search by Topic: Condition - Kidney Cancer	747
Advance search: Condition - renal cancer AND Interventions - Pazopanib	9
Advance search: Condition - renal cancer AND Interventions - Bevacizumab	53
Advance search: Condition - renal cancer AND Interventions - Sunitinib	93
Advance search: Condition - renal cancer AND Interventions - Temsirolimus	16
Advance search: Condition - renal cancer AND Interventions - Interferon	69
Advance search: Condition - renal cancer AND Interventions - interleukin	65
Advance search: Condition - renal cancer AND Interventions - Everolimus	21
Advance search: Condition - renal cancer AND Interventions - Sorafenib	67
Advance search: Condition - renal cell carcinoma AND Interventions - Pazopanib	9
Advance search: Condition - renal cell carcinoma AND Interventions - Bevacizumab	52
Advance search: Condition - renal cell carcinoma AND Interventions - Sunitinib	91
Advance search: Condition - renal cell carcinoma AND Interventions - Temsirolimus	16
Advance search: Condition - renal cell carcinoma AND Interventions - Interferon	68
Advance search: Condition - renal cell carcinoma AND Interventions - interleukin	65
Advance search: Condition - renal cell carcinoma AND Interventions - Everolimus	20
Advance search: Condition - renal cell carcinoma AND Interventions - Sorafenib	66

Total retrieved: 196 (after removing duplicates and potential exclusions)

B.2 Economic systematic review

The search strategies for the economic systematic review are presented in the following sections.

B.2.1 MEDLINE and EMBASE

Date search run: 23 November 2009

Table 77: Search for randomised controlled trials – Embase.com

#	Search History	Results
1.	'economics'/de	174965
2.	'economic aspect'/de	92015
3.	'cost'/de	45254
4.	'health care cost'/de	86127
5.	'drug cost'/de	41407
6.	'hospital cost'/de	9530
7.	'socioeconomics'/de	84286
8.	'health economics'/de	29146
9.	'pharmacoeconomics'/de	1708
10.	'fee'/exp	28062
11.	'budget'/exp	13917
12.	'economic evaluation'/exp	147273
13.	'hospital finance'/de OR 'financial management'/de	87263
14.	'health care financing'/de	9708
15.	'low cost'	15803
16.	'high cost'	5526
17.	health*care NEXT/1 cost* OR 'health care' NEXT/1 cost*	146263
18.	fiscal OR funding OR financial OR finance	264093
19.	cost NEXT/1 estimate*	1248
20.	'cost variable'	32
21.	unit NEXT/1 cost*	1228
22.	economic*:ab,ti OR pharmacoeconomic*:ab,ti OR price*:ab,ti OR pricing:ab,ti	144464
23.	(cost* NEAR/3 (treat* OR therap*)):ab,ti	19388
24.	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23	769525
25.	'pazopanib'/de OR 'sunitinib'/de OR 'sorafenib'/de OR 'bevacizumab'/de OR 'temsirolimus'/de OR 'everolimus'/de OR 'interleukin 2'/de OR 'alpha interferon'/de	94250
26.	'alpha-interferon':ab,ti OR alfaferone:ab,ti OR alferon:ab,ti OR 'alpha ferone':ab,ti OR cilferon:ab,ti OR ginterferon:ab,ti OR 'interferon-alpha':ab,ti OR introma:ab,ti OR kemron:ab,ti OR leukinferon:ab,ti OR leukinferron:ab,ti OR 'leukocyte interferon':ab,ti OR 'refecon a':ab,ti OR 'referon a3':ab,ti OR sumiferon:ab,ti OR sumipheron:ab,ti OR veldona:ab,ti	10766
27.	'biotest':ab,ti OR bioleukin:ab,ti OR 'interleukin-ii':ab,ti OR 'interleukin-2':ab,ti OR 'il-2':ab,ti OR il2:ab,ti OR 'ro-236019':ab,ti OR tcgf:ab,ti OR tsf:ab,ti	56840
28.	everolimus:ab,ti OR afinitor:ab,ti OR certican:ab,ti OR 'nvp-rad-001':ab,ti OR 'rad-001':ab,ti OR 'rad 001a':ab,ti OR rad001:ab,ti OR rad001a:ab,ti OR 'sdz rad':ab,ti	853
29.	temsirolimus:ab,ti OR 'cci-779':ab,ti OR 'cell-cycle-inhibitor-779':ab,ti OR 'nsc 683864':ab,ti OR nsc683864:ab,ti OR torisel:ab,ti	402
30.	bevacizumab:ab,ti OR avastin:ab,ti OR 'nsc 704865':ab,ti OR nsc704865:ab,ti OR 'anti-vegf':ab,ti OR 'rhumab-vegf':ab,ti	4100
31.	'bay 43-9006':ab,ti OR 'bay 439006':ab,ti OR 'bay43-9006':ab,ti OR bay439006:ab,ti OR nexavar:ab,ti OR sorafenib:ab,ti	996
32.	sunitinib:ab,ti OR sutent:ab,ti OR 'pha 2909040ad':ab,ti OR 'pha2909040ad':ab,ti OR 'su 010398':ab,ti OR 'su 011248':ab,ti OR 'su 10398':ab,ti OR su10398:ab,ti OR 'su 11248':ab,ti OR su010398:ab,ti OR 'su011248':ab,ti OR su11248:ab,ti	960
33.	armala:ab,ti OR pazopanib:ab,ti OR gw786034*:ab,ti OR (gw NEXT/1 786034*):ab,ti OR	39

#	Search History	Results
	(sb NEXT/1 710468*):ab,ti OR sb710468*:ab,ti	
34.	#25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33	121986
35.	'kidney carcinoma'/de	27437
36.	'kidney tumour'/exp	64633
37.	renal*:ab,ti OR kidney*:ab,ti OR grawit*:ab,ti OR hypernephroid*:ab,ti OR nephroid*:ab,ti	602660
38.	carcinoma*:ab,ti OR cancer*:ab,ti OR neoplasm*:ab,ti OR adeno*:ab,ti OR tumo?r*:ab,ti OR pyelocarcinoma*:ab,ti OR metastas?s:ab,ti OR oncocytoma:ab,ti	1586263
39.	#37 AND #38	68847
40.	(metanephric NEAR/2 adeno*):ab,ti	136
41.	rcc:ab,ti OR mrcc:ab,ti OR 'm-rcc':ab,ti	5722
42.	'hypernephroma':ab,ti	1196
43.	#35 OR #36 OR #39 OR #40 OR #41 OR #42	100246
44.	#24 AND #34 AND #43	192
45.	**#24 AND #34 AND #43 AND [1980-2010]/py**	**192**

B.2.2 Cochrane

Date search run: 23 November 2009

Table 78: Cochrane library search strategy

ID	Search History	Results
#1	MeSH descriptor Interferon-alpha explode all trees	2099
#2	MeSH descriptor Interleukin-2 explode all trees	702
#3	("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):ab,ti,kw	3001
#4	(biotest OR bioleukin OR "interleukin-ii" OR "interleukin-2" OR "il-2" OR il2 OR "ro-236019" OR tcgf OR tsf):ab,ti,kw	1902
#5	(everolimus OR afinitor OR certican OR "nvp-rad-001" OR "rad-001" OR "rad 001a" OR rad001 OR rad001a OR "sdz rad"):ab,ti,kw	154
#6	(temsirolimus OR "cci-779" OR "cell-cycle-inhibitor-779" OR "nsc 683864" OR nsc683864 OR torisel):ab,ti,kw	25
#7	(bevacizumab OR avastin OR "nsc 704865" OR nsc704865 OR "anti-vegf" OR "rhumab-vegf"):ab,ti,kw	236
#8	("bay 43-9006" OR "bay 439006" OR "bay43-9006" OR bay439006 OR nexavar OR sorafenib):ab,ti,kw	63
#9	(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):ab,ti,kw	37
#10	(armala OR pazopanib OR gw786034* OR sb710468* OR votrient):ab,ti,kw	2
#11	(#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10)	5875
#12	MeSH descriptor Carcinoma, Renal Cell explode all trees	301
#13	(renal* OR kidney* OR grawit* OR hypernephroid* OR nephroid*):ab,ti,kw	24198
#14	(carcinoma* OR cancer* OR neoplasm* OR adeno* OR tumo?r* OR pyelocarcinoma* OR metastas?s OR oncocytoma):ab,ti,kw	60577

ID	Search History	Results
#15	(#13 AND #14)	1811
#16	(metanephric adj2 adeno*):ab,ti,kw	0
#17	(rcc OR mrcc OR "m-rcc"):ab,ti,kw	168
#18	hypernephroma:ab,ti,kw	4
#19	(#12 OR #15 OR #16 OR #17 OR #18)	1829
#20	(#11 AND #19)	334
#21	**(#14 AND #22), from 1980 to 2009 [Technology assessments, Economic evaluations]**	**14**

Appendix C Extraction grid

C.1 Clinical systematic review

CITATION ID		**AUTHOR, YEAR**	
REVIEWER			
Characteristics of study			
Publication type	Journal/Conference abstract		
Blinding	Open-label/Single blind/Double/Triple		
Control	Active/placebo/dose ranging		
Cross-over permitted	Yes/No/Unclear		
Centre	Mutlicentre/Multicentre international/Single centre		
Phase of RCT	I/II/III/IV/Mixed/Unclear		
Efficacy analysis type	ITT/mITT/PP		
Efficacy analysis type	ITT/mITT/PP		
Country	Country(ies) study conducted in		
Study methods	Description of study methods **If patients were allowed to cross over from treatment, please record when this cross-over was permitted, and what criteria allowed them to cross-over.**		
Study duration	Length of study		
Study outcomes	List of study outcomes measured		
Trial population			
Number of patients	Number of patients screened, randomised and completed		
Patient group	Treatment naïve/cytokine pre-treated/Mixed with subgroup analysis of two populations/Mixed with no subgroup analysis		
Inclusion criteria/Exclusion criteria	Main inclusion and exclusion criteria		
Prior nephrectomy	Details of prior nephrectomy – when did patients have this, were they permitted to have this before enrolment?		
Other study details			
Study objective			
Study conclusion and comments	Authors conclusions and further comments on the study		
Statistical methodology			
Critical appraisal			

Was randomisation carried out appropriately?	
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	
Is there any evidence to suggest that the authors measured more outcomes than they reported?	
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	

Interventions

	Intervention 1	Intervention 2	Intervention 3
Interventions evaluated			
Interventions dosing			
Route of administration			
Number of patients in treatment arm			
Patient characteristics			
Age (mean/median/SD)			
Disease duration (mean/median/SD)			
% Male			
Performance status – ECOG Number of patients in each performance status per treatment arm			
Performance status – KPS Number of patients in each performance status per treatment arm			
Prior treatment 1. **Nephrectomy** 2. **Radiation/Radiotherapy** 3. **Chemotherapy naïve** Number of patients previously treated with each modality.			
Histology 1. **Clear cell** 2. **Papillary** 3. **Sarcomatoid** 4. **Granular** 5. **Any subtype other than clear cell** Number of patients previously treated with each histology.			
Down dosing Number of patients who reduced the dose per treatment arm			
Dose interruption Number of patients who had a dose reduction in each treatment arm			
Number of metastasis sites 1. **1 site** 2. **2 sites** 3. **3 or more sites** Number of patients with each number of metastatic sites.			
MSKCC prognostic factors 1. **Favourable**			

2. Intermediate **3. Poor/Low** Number of patients in each MSKCC group.			
Number of prognostic factors **1. 3 or more** **2. Less than 3** Number of patients with each number of prognostic factors.			
Dichotomous outcomes – number of patients with each outcome per treatment arm			
Response rate **1. Overall response** **2. Complete response** **3. Partial response** **4. No response** **5. Stable disease**			
Overall survival **1. OS at 1 year** **2. OS at 2 year** **3. OS at endpoint (other time point)**			
Progression free survival **1. PFS at 1 year** **2. PFS at 2 year** **3. PFS at endpoint (other time point)**			
Continuous Outcomes – mean, median, SD, SE, for each outcome			
Overall survival (ITT)			
Overall survival (Cross-over adjusted)			
Progression free survival			
Time to progression			
Time to response			
Duration of response			
FKSI-DRS Index **1. Baseline** **2. Endpoint** **3. Change from baseline**			
FKSI-15 Index **1. Baseline** **2. Endpoint** **3. Change from baseline**			
FACT-G **1. Baseline** **2. Endpoint** **3. Change from baseline**			
EQ-5D **1. Baseline** **2. Endpoint** **3. Change from baseline**			
EQ-VAS **1. Baseline** **2. Endpoint** **3. Change from baseline**			
EORTC QLQ-C30 **1. Baseline** **2. Endpoint** **3. Change from baseline**			
Treatment discontinuations - number of patients withdrawing due to each reason per treatment arm			
All withdrawals			
Withdrawals due to death			
Withdrawals due to AEs			

AEs - number of patients experiencing each AE per treatment arm			
Any AE (all grades)			
Any grade 3 or 4 AE			
Any serious AEs			
Any treatment related AEs			
Abdominal pain			
Alopecia			
ALT increased			
Altered taste			
Anaemia			
Anorexia			
Arthralgia			
AST increased			
Asthenia			
Congestive heart failure			
Depression			
Diarrhoea			
Dyspepsia			
Epistaxis			
Fatigue			
Fever			
Flank pain			
Flu-like symptoms			
Hair colour change			
Hand-foot syndrome			
Headache			
Hyperglycaemia			
Hypertension			
Hypoglycaemia			
Hypophosphataemia			
Infection			
Leucopenia			
Lymphocytopaenia			
Mucositis/stomatitis			
Nausea			
Neutropaenia			
Rash			
Skin discolouration			
Thrombocytopenia			
Total bilirubin increased			

Vomiting			

C.2 Economic systematic review

Citation ID:		**Date:**	
Study:		**Reviewer:**	
Characteristics of study			
Study objective			
Country			
Study size			
Funded by			
Evaluation scope:			
Patient population	Age		
	Gender		
	Cancer stage		
	Prognosis		
	MSKCC score		
	Line of therapy		
	Histological subtype		
	Down dosing details		
	Other		
Interventions evaluated	Intervention	Dosing detail	
	<add row if require>		
Evaluation framework:			
Type of economic evaluation			
Perspective			
Method of analysis			
Model description			
Timeframe			
Discounting			
Cost and currency			
Clinical outcomes:			
Measure of outcome			
Data source			
Utilities	Were utilities used in the analysis?		
	Derivation/estimation		
Economic outcomes:			
Resource item			
Source of resource			
Source of unit cost			
Key results:			
Effectiveness and utilities			
Resources			
Costs			
ICERs/ICs			
Subgroup analysis			
Sensitivity	Method used		
	Variables considered		
	Key results		
Authors conclusion			

Critical appraisal – Drummond Checklist		
	Y/N/U	**Commentary**
1. Was a well-defined question posed in answerable form?		
1.1. Did the study examine both costs and effects of the service(s) or programme(s)?		
1.2. Did the study involve a comparison of alternatives?		
1.3. Was a viewpoint for the analysis stated and was the study placed in any particular decision-making context?		
2. Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often)?		
2.1. Were there any important alternatives omitted?		
2.2. Was (should) a do-nothing alternative be considered?		
3. Was the effectiveness of the programme or services established?		
3.1. Was this done through a randomised, controlled clinical trial? If so, did the trial protocol reflect what would happen in regular practice?		
3.2. Was effectiveness established through an overview of clinical studies?		
3.3. Were observational data or assumptions used to establish effectiveness? If so, what are the potential biases in results?		
4. Were all the important and relevant costs and consequences for each alternative identified?		
4.1. Was the range wide enough for the research question at hand?		
4.2. Did it cover all relevant viewpoints? (Possible viewpoints include the community or social viewpoint, and those of patients and third-party payers. Other viewpoints may also be relevant depending upon the particular analysis.)		
4.3. Were the capital costs, as well as operating costs, included?		
5. Were costs and consequences measured accurately in appropriate physical units (e.g. hours of nursing time, number of physician visits, lost work-days, gained life years)?		
5.1. Were any of the identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?		
5.2. Were there any special circumstances (e.g., joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?		
6. Were the cost and consequences valued credibly?		
6.1. Were the sources of all values clearly identified? (Possible sources include market values, patient or client preferences and views, policy-makers' views and health professionals' judgements)		
6.2. Were market values employed for changes involving resources gained or depleted?		
6.3. Where market values were absent (e.g. volunteer labour), or market values did not reflect actual values (such as clinic space donated at a reduced rate), were adjustments made to approximate market values?		
6.4. Was the valuation of consequences appropriate for the question posed (i.e. has the appropriate type or types of analysis – cost-effectiveness, cost-benefit, cost-utility – been selected)?		
7. Were costs and consequences adjusted for differential timing?		
7.1. Were costs and consequences that occur in the future 'discounted' to their present values?		
7.2. Was there any justification given for the discount rate used?		
8. Was an incremental analysis of costs and consequences of alternatives performed?		

8.1. Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits, or utilities generated?			
9. Was allowance made for uncertainty in the estimates of costs and consequences?			
9.1. If data on costs and consequences were stochastic (randomly determined sequence of observations), were appropriate statistical analyses performed?			
9.2. If a sensitivity analysis was employed, was justification provided for the range of values (or for key study parameters)?			
9.3. Were the study results sensitive to changes in the values (within the assumed range for sensitivity analysis, or within the confidence interval around the ratio of costs to consequences)?			
10. Did the presentation and discussion of study results include all issues of concern to users?			
10.1. Were the conclusions of the analysis based on some overall index or ratio of costs to consequences (e.g. cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?			
10.2. Were the results compared with those of others who have investigated the same question? If so, were allowances made for potential differences in study methodology?			
10.3. Did the study discuss the generalisability of the results to other settings and patient/client groups?			
10.4. Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and consequences, or relevant ethical issues)?			
10.5. Did the study discuss issues of implementation, such as the feasibility of adopting the 'preferred' programme given existing financial or other constraints, and whether any freed resources could be redeployed to other worthwhile programmes?			
Critical appraisal - Philips et al 2006. These refer specifically to decision-analytic models and should only be answers if the paper includes a model.			
Section		**Y/N/U**	**Commentary**
Method of analysis	Is the pre-evaluation data analysis methodology based on justifiable statistical and epidemiological techniques?		
	Has the evidence regarding the model structure been described? Is the structure of the model consistent with a coherent theory of the health condition under evaluation?		
	Are the sources of data used to develop the structure of the model specified?		
	Are the causal relationships described by the model structure justified appropriately?		
	Are the structural assumptions transparent and justified?		
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the evaluation?		
	Do the disease states or the care pathways reflect the underlying biological process of the disease in question and the impact of interventions?		
	Is the cycle length defined and justified in terms of the natural history of disease?		
	Are transition probabilities calculated appropriately?		
	Has a half cycle correction been applied to both cost and outcome? If not, has this omission been justified?		
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?		
Sensitivity analysis	Have alternative extrapolation assumptions been explored through sensitivity analysis?		

	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?		
	Have the four principal types of uncertainty been addressed (methodological, structural, heterogeneity, parameters)? If not, has the omission of particular forms of uncertainty been justified?		
	Have methodological uncertainties been addressed by running alternative versions of the evaluation with different methodological assumptions?		
Author conclusions	Is there evidence that the mathematical logic of the evaluation has been tested thoroughly before use?		
	If the evaluation has been calibrated against independent data, have any differences been explained and justified?		

Appendix D Quality assessment of clinical studies

Study name	Jadad score	Allocation grade	Randomisation	Baseline characteristics	Blinding	Withdrawal	Outcome selection and reporting	Statistical analysis
VEG105192	5	A	Yes. Patients were centrally randomly assigned in a 2:1 ratio to pazopanib or placebo. Eligible patients were stratified on the basis of baseline ECOG PS (0 vs. 1), prior nephrectomy (yes vs. no), and prior systemic therapy for advanced RCC (cytokine-pretreated vs.treatment naïve) and were randomised by GSK Biomedical Data Sciences Department using GSK interactive voice response system (IVRS) called RAMOS (Registration And Medication Ordering System).	Yes. Baseline comparability was achieved between the two groups in terms of age, gender, race, histology, disease duration, organs involved, ECOG performance status and MSKCC risk category.	Yes. Adequate blinding was achieved by using matching placebo tablets. Additionally, disease assessments were conducted by independent reviewers who were also blinded to treatment assignment.	No. Reasons for withdrawal of patients were reported adequately. Patients mainly withdrew due to the following reasons: disease progression; death; AEAEs; lost to follow-up; protocol violation; patient or investigator's decision; or other reasons.	No. The authors reported all the outcomes as specified in the protocol of the study.	Yes. An ITT analysis was used for efficacy evaluation and appropriate methods were used to account for missing data.
Motzer 2009	2	B	Not Clear. Patients were randomised using permuted block design. The method of allocation concealment was not reported in the study.	Yes. The two treatment groups were similar in terms of their performance status, prior chemotherapy and histology.	Yes. An independent central review committee evaluating the radiographs was blinded to the treatment allocation. Blinding status of investigators was unclear.	No. There were no unexpected imbalances in the drop-outs between the groups. Withdrawals and reasons for all cause withdrawals were reported in the study.	Not clear.	Yes. The primary end point was analysed in all patients assigned to a study group, according to the intention-to-treat principle. Safety analyses were performed on the basis of the treatment actually received. Method of handling

Study name	Jadad score	Allocation grade	Randomisation	Baseline characteristics	Blinding	Withdrawal	Outcome selection and reporting	Statistical analysis
								missing data was not reported in the study.
Escudier 2009	2	B	Not clear. Patients were randomly assigned (1:1) to sorafenib or interferon and were stratified by MSKCC classification and region.	Yes. Baseline characteristics were similar between groups in terms of histology, performance and prognostic factors.	No. This was an open-label study. However, data from independent blinded radiologic review were the primary data for determination of radiologic progression of period I.	No. Treatment discontinuations due to AEs and death were reported for both the arms.	No. The authors measured all the outcomes that were reported.	Yes. For efficacy analysis, ITT population was used. Safety population was mITT. Appropriate statistical methods were used in the study.
Target Study	3	B	Not clear. Patients were stratified according to country and MSKCC prognostic score and randomly assigned to study groups in a 1:1 ratio with a block size of four.	Yes. Baseline characteristics were comparable between study groups in terms of age, weight, ECOG score etc.	Yes. This was a triple-blind study. Investigators and independent radiologists who were unaware of the study-group assignments assessed study outcomes.	No. The reasons for withdrawals were reported in the study.	Not clear	Yes. ITT and mITT approaches were used to analyse efficacy and safety data, respectively. Details of handling missing data were not reported.
AVOREN trial	4	A	Yes. Randomisation was done centrally with a block design procedure and stratified according to country and MSKCC risk group. Patient randomisation list was kept in secure location and was not available to any person directly involved in the study other than the interactive voice recognition system provider and the randomisation manager at Roche.	Yes. Authors reported that the arms were balanced with regard to baseline disease and demographic characteristics.	Yes. This was a double blind study. The method of blinding was unclear.	Yes. Reasons for withdrawals and all cause withdrawals were reported in the study.	Not clear	Yes. Primary efficacy analysis was done by intention-to treat approach. For secondary efficacy analysis patients with measurable disease at baseline were included. All patients who were randomised and exposed to study medication were included in the safety analysis. For safety analysis, patients were assigned to treatment groups on the basis of what they actually received, with patients in the placebo arm receiving one or more

Study name	Jadad score	Allocation grade	Randomisation	Baseline characteristics	Blinding	Withdrawal	Outcome selection and reporting	Statistical analysis
								doses of bevacizumab being assigned to the bevacizumab arm.
CALGB 90206	2	B	Not clear. Patients were randomised according to stratified random block design. Patients were stratified by nephrectomy status and number of adverse prognostic factors. The method of allocation concealment was not reported in the study.	Yes. The two treatment groups were similar in terms of their performance status, prior chemotherapy and histology.	No. This was an open label trial.	No. There were no unexpected imbalances in the drop-outs between the groups.	Unclear. It was unclear whether the authors measured more outcomes than they reported.	Yes. Patients who discontinued treatment for reasons other than progression were observed for disease progression or death. An intention-to-treat approach was used in the analysis.
Global ARCC trial	2	B	Not Clear. Patients were randomly assigned in equal proportions, with the use of permuted blocks of three, to one of three treatment groups. Method of concealment of allocation was not reported in the study.	Yes; The three treatment groups were well balanced on the basis of age, sex, and performance-status score.	No. This was an open-label trial.	No; Reasons for treatment discontinuation included disease progression, AEs, symptomatic deterioration, death, patient request, other and protocol violation. A total of 19 patients were lost to follow-up.	No; the authors reported measured outcomes only.	Yes; The primary end point was calculated on an intention-to-treat basis. An appropriate statistical analysis was used. Details regarding handling of missing data were not reported.
Negrier 2007	1	A	Not clear. Randomization was stratified by participating centre by using a block method with a block size of 4, and it was performed centrally through a specific website.	Yes. Author has reported that only few significant differences were detected between comparison groups. More non-IFN-treated than IFN-treated patients had abdominal lymph nodes (31.2% vs. 21.8%; P = .02), and less non-IFN-treated than IFN-treated patients had elevated serum lactate dehydrogenase (LDH) levels (16.2% vs.	No. Treatments were administered unblinded.	Not clear. Details regarding withdrawals were not reported.	No. Four additional per-protocol analyses were performed: 1) after exclusion of the 18 patients with major protocol deviation, 2) after exclusion of the 58 patients crossed over to another treatment, 3) on the 386 patients with proven clear cell renal cancer, or 4) on the 270 patients not receiving second-line treatment. All yielded	Yes. Efficacy analysis was done on ITT basis. Method of handling missing data was not reported in the study.

Study name	Jadad score	Allocation grade	Randomisation	Baseline characteristics	Blinding	Withdrawal	Outcome selection and reporting	Statistical analysis
				25.1%; P = .03). Among non-IL-2-treated patients, 73.1% had normal hemoglobin levels versus 63.8% among patients receiving IL-2 (P = .03). Comparison groups were overall considered well balanced.			results similar to those of the first analysis.	
CRECY Trial	2	A	Yes. Randomisation was performed centrally by an interactive computerised procedure at the study data-monitoring centre. Randomisation was stratified according to centre.	Yes. Authors stated that there were no significant differences in patient characteristics among three treatment groups.	Yes. Blinded external committee reviewed treatment response.	Not clear. Number of patients and reasons for withdrawal were not reported.	Not clear	Yes. An ITT analyses was performed. Method of handling of missing data was not reported.
MRC RE01	1	A	Not clear. A minimisation method was used and patients were stratified by centre, nephrectomy and by whether there were single or multiple metastases. Concealment of allocation was adequate. Randomisation was by telephone call to the MRC Cancer Trials Office.	Yes. Authors stated that characteristics of patients were similar in both treatment groups.	Not clear. It was unclear whether study was blinded or not.	Not clear. The details regarding withdrawal were not reported in the study. Protocol deviations were reported by the author.	Not clear	Yes. An ITT analyses was performed for primary efficacy outcome. Method of handling of missing data was not reported.
Steineck 1990	1	B	Not clear. Method of randomisation and allocation concealment was not reported in the study.	Yes. Baseline characteristics seem to be comparable in terms of age, gender and other demographic	Yes. The outcome assessor (radiologist) was blinded to the treatment. Blinding status of patients and	No. Withdrawals and reasons for all cause withdrawals were not reported in the study.	Unclear. It was unclear whether the authors measured more outcomes than they reported.	Yes. An ITT analysis for efficacy and safety was carried out. For more strict evaluation of efficacy nine patients

Study name	Jadad score	Allocation grade	Randomisation	Baseline characteristics	Blinding	Withdrawal	Outcome selection and reporting	Statistical analysis
				characteristics.	investigators was unclear.			were excluded from the analysis. The exclusion of these patients did not change the proportion of responding patients.
Kriegmair 1995	1	B	Not clear. Patients were randomised in blocks of five to each treatment group. Method of concealment of allocation was unclear.	Yes. Baseline comparability was achieved between the two groups in terms of age, gender, performance status and distribution of the tumour lesions.	Not clear. Blinding of patients, investigators, statistician or outcome assessor was not reported.	Yes. There were unexpected imbalances in the drop-outs between the groups. Three patients in IFN plus vinblastine group and 10 patients in medroxyprogestrone group withdrew the informed consent.	Not clear. It is unclear whether authors measured more outcomes than they reported.	No. A PP analysis was used for efficacy and safety evaluations. Method for handling missing data was not reported.
Pyrhonen 1999	2	B	Not clear. The method of randomisation and allocation concealment was not reported.	Yes. The treatment groups were well balanced for all measured baseline demographic and disease characteristics.	Not clear. Films of patients with objective response were reviewed by a single central radiologist and the principal investigators from the two centres not treating the patient.	No. In this study, no patients were lost to follow-up at the time of this report, and follow-up of all surviving patients is continuing. The reasons for withdrawals were reported adequately.	Not clear	Yes. Data was analysed using an ITT analyses.

Appendix E Critical appraisal of economic studies

Study:	**Mickisch 2009**	
Critical appraisal – Drummond Checklist		
	Y/N/U	**Commentary**
1. Was a well-defined question posed in answerable form?		
1.1. Did the study examine both costs and effects of the service(s) or programme(s)?	Y	The study examined both costs and effects of the treatments.
1.2. Did the study involve a comparison of alternatives?	Y	The study compared sunitinib with combination of bevacizumab and IFN-alfa.
1.3. Was a viewpoint for the analysis stated and was the study placed in any particular decision-making context?	Y	The viewpoint of the study was clearly stated.
2. Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often)?		
2.1. Were there any important alternatives omitted?	Y	There are many alternatives used in the treatment of disease in question. Only two of those were included in this analysis.
2.2. Was (should) a do-nothing alternative be considered?	N	Ethically, do-nothing alternative should not be used in cancer trials.
3. Was the effectiveness of the programme or services established?		
3.1. Was this done through a randomised, controlled clinical trial? If so, did the trial protocol reflect what would happen in regular practice?	N	The data was obtained from clinical trials.
3.2. Was effectiveness established through an overview of clinical studies?	Y	The data was obtained from clinical trials.
3.3. Were observational data or assumptions used to establish effectiveness? If so, what are the potential biases in results?	N	Observational data was not used to establish the effectiveness of the study.
4. Were all the important and relevant costs and consequences for each alternative identified?		
4.1. Was the range wide enough for the research question at hand?	N	Only the management cost for AE were considered
4.2. Did it cover all relevant viewpoints? (Possible viewpoints include the community or social viewpoint, and those of patients and third-party payers. Other viewpoints may also be relevant depending upon the particular analysis.)	N	Provider's viewpoint was considered
4.3. Were the capital costs, as well as operating costs, included?	N	Only the management cost for AE were considered
5. Were costs and consequences measured accurately in appropriate physical units (e.g. hours of nursing time, number of physician visits, lost work-days, gained life years)?		
5.1. Were any of the identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?	N	Costs and consequences were measured accurately and in appropriate physical units.
5.2. Were there any special circumstances (e.g., joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?	N	No such circumstances were discussed.
6. Were the cost and consequences valued credibly?		
6.1. Were the sources of all values clearly identified? (Possible sources include market values, patient or client preferences and views, policy-makers' views and health professionals' judgements)	Y	All important sources were clearly identified.
6.2. Were market values employed for changes involving resources gained or depleted?	N	Not reported
6.3. Where market values were absent (e.g. volunteer labour), or market values did not reflect actual values (such as clinic space donated at a reduced rate), were adjustments made to approximate market values?	N	Not reported
6.4. Was the valuation of consequences appropriate for the question posed (i.e. has the appropriate type or types of analysis – cost-effectiveness, cost-benefit, cost-utility – been selected)?	Y	The type of costs analysis was appropriate to answer the study question.
7. Were costs and consequences adjusted for differential timing?		
7.1. Were costs and consequences that occur in the future 'discounted' to their present values?	N	Not reported
7.2. Was there any justification given for the discount rate used?	N	Not reported

8. Was an incremental analysis of costs and consequences of alternatives performed?		
8.1. Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits, or utilities generated?	N	This was a simple costs analysis study. Incremental costs analysis was not done.
9. Was allowance made for uncertainty in the estimates of costs and consequences?		
9.1. If data on costs and consequences were stochastic (randomly determined sequence of observations), were appropriate statistical analyses performed?	Y	Appropriate statistical analyses were performed.
9.2. If a sensitivity analysis was employed, was justification provided for the range of values (or for key study parameters)?	N	Details of sensitivity analyses were not reported.
9.3. Were the study results sensitive to changes in the values (within the assumed range for sensitivity analysis, or within the confidence interval around the ratio of costs to consequences)?	N	The study results were not sensitive to changes in various parameters. It proved that analytic model was robust.
10. Did the presentation and discussion of study results include all issues of concern to users?		
10.1. Were the conclusions of the analysis based on some overall index or ratio of costs to consequences (e.g. cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?	N	It was a cost study.
10.2. Were the results compared with those of others who have investigated the same question? If so, were allowances made for potential differences in study methodology?	N	Authors compared the results with those of others who investigated the same question.
10.3. Did the study discuss the generalisability of the results to other settings and patient/client groups?	N	Generalisability of the results was not discussed.
10.4. Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and consequences, or relevant ethical issues)?	N	Not reported
10.5. Did the study discuss issues of implementation, such as the feasibility of adopting the 'preferred' programme given existing financial or other constraints, and whether any freed resources could be redeployed to other worthwhile programmes?	N	No such discussion was reported.

Critical appraisal - Philips et al 2006

Section		Y/N/U	Commentary
Method of analysis	Is the pre-evaluation data analysis methodology based on justifiable statistical and epidemiological techniques?	N	Not reported
	Has the evidence regarding the model structure been described? Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	N	Ns such evidence was described in the study.
	Are the sources of data used to develop the structure of the model specified?	Y	All sources of data were specified.
	Are the causal relationships described by the model structure justified appropriately?	N	It was justified.
	Are the structural assumptions transparent and justified?	N	Structural assumptions were transparent.
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the evaluation?	N	Structural assumptions were reasonable given the objective and perspective.
	Do the disease states or the care pathways reflect the underlying biological process of the disease in question and the impact of interventions?	N	Not reported
	Is the cycle length defined and justified in terms of the natural history of disease?	N	Not reported
	Are transition probabilities calculated appropriately?	N	Not reported
	Has a half cycle correction been applied to both cost and outcome? If not, has this omission been justified?	N	Not reported
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	N	No such assumptions were reported.
Sensitivity analysis	Have alternative extrapolation assumptions been explored through sensitivity analysis?	Y	All possible alternatives were explored through sensitivity analysis.

	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	N	Not reported
	Have the four principal types of uncertainty been addressed (methodological, structural, heterogeneity, parameters)? If not, has the omission of particular forms of uncertainty been justified?	N	Detail of methodology of sensitivity analysis was not reported.
	Have methodological uncertainties been addressed by running alternative versions of the evaluation with different methodological assumptions?	Y	Modification of basic clinical and economic assumptions (hospitalisation costs and the main cost-driving AEs) showed that the model remained stable over the entire range of plausible values for a given parameter
Author conclusions	Is there evidence that the mathematical logic of the evaluation has been tested thoroughly before use?	N	No such evidence was reported.
	If the evaluation has been calibrated against independent data, have any differences been explained and justified?	N	It was not reported.

Study:	**Remak 2008**	
Critical appraisal – Drummond Checklist		
	Y/N/U	**Commentary**
1. Was a well-defined question posed in answerable form?		
1.1. Did the study examine both costs and effects of the service(s) or programme(s)?	Y	The study examined both costs and effects of treatments.
1.2. Did the study involve a comparison of alternatives?	Y	The study compared sunitinib with IL-2 and IFN.
1.3. Was a viewpoint for the analysis stated and was the study placed in any particular decision-making context?	Y	View point of analysis was clearly identified as US societal perspective.
2. Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often)?		
2.1. Were there any important alternatives omitted?	N	It seems that important alternatives were used for comparison.
2.2. Was (should) a do-nothing alternative be considered?	U	Comparison with observation or best supportive care could have been performed.
3. Was the effectiveness of the programme or services established?		
3.1. Was this done through a randomised, controlled clinical trial? If so, did the trial protocol reflect what would happen in regular practice?	U	Effectiveness of intervention was established through randomised controlled trial which may not adequately reflect routine clinical practice.
3.2. Was effectiveness established through an overview of clinical studies?	N	Data were derived from an RCT.
3.3. Were observational data or assumptions used to establish effectiveness? If so, what are the potential biases in results?	N	Data were derived from an RCT.
4. Were all the important and relevant costs and consequences for each alternative identified?		
4.1. Was the range wide enough for the research question at hand?	N	The study adopted societal perspective however indirect cost were not included.
4.2. Did it cover all relevant viewpoints? (Possible viewpoints include the community or social viewpoint, and those of patients and third-party payers. Other viewpoints may also be relevant depending upon the particular analysis.)	N	A narrow perspective was used and only direct medical costs were included. Burden of disease on family and care givers and indirect cost to society could be considered.
4.3. Were the capital costs, as well as operating costs, included?	U	Intervention is unlikely to introduce capital expenditure.
5. Were costs and consequences measured accurately in appropriate physical units (e.g. hours of nursing time, number of physician visits, lost work-days, gained life years)?		
5.1. Were any of the identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?	N	All identified items were measured and included in analysis.
5.2. Were there any special circumstances (e.g., joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?	U	Authors do not discuss about any circumstances which made measurement difficult.
6. Were the cost and consequences valued credibly?		

6.1. Were the sources of all values clearly identified? (Possible sources include market values, patient or client preferences and views, policy-makers' views and health professionals' judgements)	Y	Sources of all values were clearly reported.
6.2. Were market values employed for changes involving resources gained or depleted?	Y	Market values were used for resource use and their source was clearly reported.
6.3. Where market values were absent (e.g. volunteer labour), or market values did not reflect actual values (such as clinic space donated at a reduced rate), were adjustments made to approximate market values?	U	Authors do not report any circumstances were market values were absent.
6.4. Was the valuation of consequences appropriate for the question posed (i.e. has the appropriate type or types of analysis – cost-effectiveness, cost-benefit, cost-utility – been selected)?	Y	Cost-effectiveness and cost-utility analysis was performed.
7. Were costs and consequences adjusted for differential timing?		
7.1. Were costs and consequences that occur in the future 'discounted' to their present values?	Y	All costs and outcomes were discounted at 5% annually.
7.2. Was there any justification given for the discount rate used?	N	No justification provided for the discount rate used.
8. Was an incremental analysis of costs and consequences of alternatives performed?		
8.1. Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits, or utilities generated?	Y	Incremental costs were reported.
9. Was allowance made for uncertainty in the estimates of costs and consequences?		
9.1. If data on costs and consequences were stochastic (randomly determined sequence of observations), were appropriate statistical analyses performed?	Y	Statistical analyses performed were appropriate.
9.2. If a sensitivity analysis was employed, was justification provided for the range of values (or for key study parameters)?	Y	Authors stated that one-way deterministic sensitivity analysis was conducted using extreme values (reference case estimate ± 20%).
9.3. Were the study results sensitive to changes in the values (within the assumed range for sensitivity analysis, or within the confidence interval around the ratio of costs to consequences)?	Y	Deterministic sensitivity analyses showed the results to be sensitive to the utility values during treatment, costs of Sunitinib and cost of BSC.
10. Did the presentation and discussion of study results include all issues of concern to users?		
10.1. Were the conclusions of the analysis based on some overall index or ratio of costs to consequences (e.g. cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?	Y	Conclusion of analysis was based on ICER and ICUR.
10.2. Were the results compared with those of others who have investigated the same question? If so, were allowances made for potential differences in study methodology?	N	Authors do not discuss the results in comparison with other studies.
10.3. Did the study discuss the generalisability of the results to other settings and patient/client groups?	Y	Authors acknowledge that the use of clinical trial data is the major study limitation as it may not adequately reflect routine clinical practice.
10.4. Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and consequences, or relevant ethical issues)?	N	Not discussed in detail however authors discussed about threshold limit for acceptance of cost effectiveness.
10.5. Did the study discuss issues of implementation, such as the feasibility of adopting the 'preferred' programme given existing financial or other constraints, and whether any freed resources could be redeployed to other worthwhile programmes?	N	The issue was not discussed.

Critical appraisal - Philips et al 2006

Section		Y/N/U	Commentary
Method of analysis	Is the pre-evaluation data analysis methodology based on justifiable statistical and epidemiological techniques?	Y	Data analysis methodology is justifiable.
	Has the evidence regarding the model structure been described? Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y	The structure of model was well defined.
	Are the sources of data used to develop the structure of the model specified?	Y	Data sources are well reported.
	Are the causal relationships described by the model structure justified appropriately?	Y	Casual relationship described by the model structure was justified.

	Are the structural assumptions transparent and justified?	Y	Model assumptions were clearly reported.
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the evaluation?	Y	Structural assumptions seem justified.
	Do the disease states or the care pathways reflect the underlying biological process of the disease in question and the impact of interventions?	Y	Yes, disease states reflect the underlying biological process of the disease in question and the impact of interventions.
	Is the cycle length defined and justified in terms of the natural history of disease?	Y	Cycle length was defined and seems justified.
	Are transition probabilities calculated appropriately?	U	Cumulative survival probabilities were converted to 6-week cycle probabilities.
	Has a half cycle correction been applied to both cost and outcome? If not, has this omission been justified?	U	Authors do not report about half cycle correction.
	Have assumptions regarding the continuing effect of treatment once treatment is complete been documented and justified?	U	Not applicable. Patients received therapy until disease progression after which patients were switched to second line treatment or BSC.
Sensitivity analysis	Have alternative extrapolation assumptions been explored through sensitivity analysis?	N	Short term survival data were extrapolated to model long term outcome. Alternative techniques for this were not explored.
	Have alternative assumptions regarding the continuing effect of treatment been explored through sensitivity analysis?	U	Not applicable. Patients received therapy until disease progression after which patients were switched to second line treatment or BSC.
	Have the four principal types of uncertainty been addressed (methodological, structural, heterogeneity, parameters)? If not, has the omission of particular forms of uncertainty been justified?	N	Sensitivity analyses of model parameters were performed.
	Have methodological uncertainties been addressed by running alternative versions of the evaluation with different methodological assumptions?	N	Different methodological assumptions were not tested in sensitivity analysis.
Author conclusions	Is there evidence that the mathematical logic of the evaluation has been tested thoroughly before use?	Y	The study employed well established Markov modelling technique.
	If the evaluation has been calibrated against independent data, have any differences been explained and justified?	N	Authors do not discuss the results in comparison with other studies.

Appendix F Statistical methods

F.1 Pooling of effect estimates

Statistical methods for pooling effect estimates are reported in the following sections.

F.1.1 Dichotomous data

Dichotomous outcomes were summarised as the RR ratio (or odds ratio). The RR is the ratio of risks of the event in the treatment group relative to the risk of the event in the control group. The odds ratio is the ratio of odds of the event in the treatment group relative to the odds of the event in the control group. Risks and odds are defined as follows

$$Risk = \frac{n}{N} \quad (1)$$

$$Odds = \frac{n}{N-n} \quad (2)$$

where n represents the number of patients with the event and N represents the number of patients observed (generally the Intention To Treat population for that treatment group).

Meta-analysis was performed in Stata statistical software. Fixed-effects estimates were calculated according to the Mantel-Haenszel model, and random-effects estimates according to the method of DerSimonian and Laird.

F.1.2 Continuous outcomes

Continuous outcomes were summarised as the weighted mean difference. The weighted mean difference is calculated as the difference between the mean outcome values for the treatment and control groups. Meta-analysis was performed in Stata statistical software.

F.1.3 Fixed and random effects models

A pooled, meta-analysis of direct comparisons of outcomes between pairs of treatments may be conducted using fixed- or random-effects statistical techniques if studies are sufficiently homogeneous for the mathematical and statistical assumptions underlying those techniques to hold true. Ultimately this is a matter of judgement but, in order to inform such a judgement, the relevant assumptions must be understood and the factors that determine how well the assumptions hold must also be understood. Since the relevant assumptions relate to consistency in the effect estimate between studies, all sources of possible heterogeneity must be appreciated.

The key assumption underlying a fixed effects meta-analysis model is that the effect sizes measurable in each pooled study are identical. The key assumption underlying a random effects meta-analysis model is that the effect sizes measurable in each pooled study are different, but exchangeable.

There is no reason to believe the measurable effect sizes are systematically different, but that the measurable effect sizes may be considered to be drawn from some statistical distribution. In both cases, the effect size actually observed will differ from the measurable effect size, only due to sampling variation. Mechanistic methods exist for choosing between random and fixed effects models, but are not well-regarded (Egger 2001).

Adopting a random effects model in a pooled direct analysis gives more weight to smaller studies than a fixed effects model. Depending on the number of studies identified and their variation in terms of study and effect size, the decision about whether the model should have fixed or random effects may not have a large impact on pooled, direct analyses conducted. Random effects models are common in indirect analyses – this could be due to the additional heterogeneity that exists in such analyses which renders the assumptions underlying a fixed effects model less reasonable.

F.1.4 Studies with more than two treatment arms

Where studies included more than one treatment arm with the same intervention, such as two different doses of the same drug, and the analysis plan did not include a separation of these two treatment arms, data for the two arms were pooled. For dichotomous data, this simply included summing the ITT number and the number of events for each arm. For continuous data, the mean for each group was weighted according to the number of patients in each group, to calculate a weighted mean across the two groups.

F.2 Identifying sources of heterogeneity

Identification of the possible sources of heterogeneity is primarily a clinical rather than statistical matter. The terms "clinical heterogeneity" and "statistical heterogeneity" are sometimes used – the latter being the quantitative measurement, through statistical techniques, of the observed "clinical heterogeneity" (Thompson 1994).

There are some suggestions as to how it is possible to systematically review sources of heterogeneity by considering clinical and methodological diversity (Thompson 1994). Clinical diversity includes factors such as study location, setting, age, sex, diagnosis, disease severity of participants, concomitant treatments, dose or intensity of the intervention, and outcome definition. Methodological diversity includes such factors as trial design (e.g. parallel group or crossover trial), randomisation method (by cluster, individual or other), study quality (which could be measured in various ways), extent of withdrawals and analysis method (e.g. whether analyses are conducted according to the ITT principle where losses to follow-up occur). Treatment effect, whether measured in absolute or relative terms may be associated with the levels of underlying risk in patients in a trial, which may vary between trials. Attempts to account for this must be made with care (Sharp 1996).

Focus should be placed on those variables which best explain variation in outcome, whilst recognizing the dangers of over-interpretation that may exist where there are a large number of sources of possible heterogeneity relative to the number of studies included in the pooling.

F.3 Balancing multiplicity and heterogeneity

Various publications consider the conduct, reporting and mechanics of systematic reviews and meta-analyses, but few consider the specific question about what studies should be included in any such analysis (Moher 1999).

With several studies in a pooled analysis, one should aim to be able to assess both how effective the treatment is on average across the included studies, and to what extent the sources of heterogeneity explain the results obtained. With few included studies this may not be possible.

There is a balance to be struck between having few analyses focussed on primary outcomes and having analyses that contain sufficiently homogenous studies. This is because high levels of stratification will produce many pooled analyses of very similar studies, low levels of stratification will produce few pooled analyses of possibly diverse studies.

The relative advantages and disadvantages of low and high degrees of stratification are summarised in Table 79.

Table 79: Advantages and disadvantages of high and low stratification in analyses

Low stratification	High stratification
Advantages: • Few analyses • Less chance of "false positive" results • Able to focus on main issues	Advantages: • Pooled studies are very similar, so heterogeneity is likely to be less of a problem
Disadvantages: • Larger heterogeneity between studies • Results therefore less usable	Disadvantages: • Multiplicity • Greater chance of "false positive" results despite no true effect (type II error) • Large disparities in numbers of studies reporting each stratification combination • Some stratification combinations may contain no studies at all

It is tempting to use labels such as "robust", "defensible" and "poor" to describe either increasingly or decreasingly stratified pooled analyses, however, this does not capture the competing hazards of heterogeneity and multiplicity.

Stratifying analyses by all identifiable sources of heterogeneity may result in several, small analyses being conducted from which the relative efficacy of treatments may be difficult to interpret, particularly after allowing for the chance of "false positive" results. Not sufficiently stratifying the analyses may lead to the pooling of very heterogeneous studies creating pooled relative efficacy measures of limited meaning.

F.4 Publication bias

The practice of reporting certain outcomes or entire pieces of research depending on the results obtained leads to publication bias. Where publication bias exists, systematic reviews can simply be a summary of the prejudices to which the authors and researchers in that field are disposed. Systematic reviews aim to be representative of all studies, not just a subset selected according to the statistical significance or direction of their findings.

In all systematic reviews, irrespective of source funding it is important to assess as far as possible, qualitatively and statistically, the possible extent of publication bias.

Any bias is difficult to detect, though some statistical methods exist, such as examining the gradient of best fit straight lines in funnel plots, which are graphs of effect size

against the precision of such estimates (Egger 2001; Williamson 2005). Negative gradients in such plots are suggestive of greater tendencies to publish research presenting stronger positive measures of association. The Egger test has often been used, but a more reliable test by Peters et al has recently been proposed (Peters 2006).

F.5 Indirect and mixed comparisons

For clarification, Heron uses the following definitions.

- Direct comparison - an estimation of the relative treatment effect (or other relative characteristic) of one technology compared to another informed only by head-to-head RCTs of these technologies.
- Indirect comparison - an estimation of the relative treatment effect (or other relative characteristic) of one technology compared to another informed by RCTs of those technologies against a common comparator technology rather than from one or more head-to-head trials.
- Mixed treatment comparison - an estimation of the relative treatment effect (or other relative characteristic) of one technology compared to another informed from both: (i) evidence from head-to-head trials of those technologies, and (ii) trials of those technologies against a common comparator technology.

References

Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, et al. (1993) 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. 85(5): 365-376.

Alemao E, Yang S, Purvis J. (2009) Loss of work activity and productivity in caregivers attending to patients with advanced renal cell carcinoma treated with temsirolimus or interferon-alfa: evaluations from a phase 3 randomized trial. European Journal of Cancer Supplements. 7(2): 441 (Abstract P-7159).

Athar U, Gentile TC. (2008) Treatment options for metastatic renal cell carcinoma: a review. Can J Urol. 15(2): 3954-3966.

Autier J, Escudier B, Wechsler J, Spatz A, Robert C. (2008) Prospective study of the cutaneous adverse effects of sorafenib, a novel multikinase inhibitor. Arch Dermatol. 144(7): 886-892.

Bajetta E, Ravaud A, Bracarda S, Negrier S, Szczylik C, et al. (2008) Efficacy and safety of first-line bevacizumab (BEV) plus interferon-2a (IFN) in patients (pts) >65 years with metastatic renal cell carcinoma (mRCC). J Clin Oncol. 26(May 20 Supplement): Abstract 5095.

Bellmunt J, Gonzalez-Larriba JL, Climent MA, Lopez-Vivanco G, Urruticoechea L, et al. (2007) Sorafenib TARGET trial results in Spanish patients. Clin Transl Oncol. 9(10): 671-673.

Bellmunt J, Melichar B, Bracarda S, Negrier S, Ravaud A, et al. (2009) Bevacizumab (BEV) and interferon (IFN) therapy does not increase risk of cardiac events in metastatic renal cell carcinoma (mRCC). European Journal of Cancer Supplements. 7(2): 429 (Abstract P-7121).

Bellmunt J, Szczylik C, Feingold J, Strahs A, Berkenblit A. (2008) Temsirolimus safety profile and management of toxic effects in patients with advanced renal cell carcinoma and poor prognostic features. Ann Oncol. 19(8): 1387-1392.

BMJ Evidence Centre. Renal cell carcinoma. <http://bestpractice.bmj.com/best-practice/monograph/261/follow-up/prognosis.html>

Bracarda S, Bellmunt J, Negrier S, Melichar B, Ravaud A, et al. (2009) What is the impact of subsequent antineoplastic therapy on overall survival (OS) following first-line bevacizumab (BEV)/interferonalpha2a (IFN) in metastatic renal cell carcinoma (mRCC)? - Experience from AVOREN. European Journal of Cancer Supplements. 7(2): 431 (Abtract P-7126).

Bracarda S, Koralewski P, Pluzanska A, Ravaud A, Szczylik C, et al. (2007) Bevacizumab/interferon-alpha2a provides a progression-free survival benefit in all prespecified patient subgroups as first-line treatment of metastatic renal cell carcinoma (AVOREN). European Journal of Cancer Supplements. 5(4): 281 (Abstract 4008).

Bucher HC, Guyatt GH, Griffith LE, Walter SD. (1997) The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 50(6): 683-691.

Bukowski R, Cella D, Gondek K, Escudier B. (2007a) Effects of sorafenib on symptoms and quality of life: results from a large randomized placebo-controlled study in renal cancer. Am J Clin Oncol. 30(3): 220-227.

Bukowski R, Eisen T, Stadler WM, Szczylik C, Oudard S, et al. (2009) 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. European Journal of Cancer Supplements. 7(2): 432 (Abstract P-7130).

Bukowski RM, Eisen T, Szczylik C, Stadler WM, Simantov R, et al. (2007b) Final results of the randomized phase III trial of soratinib in advanced renal cell carcinoma: Survival and biomarker analysis. J Clin Oncol. 27(18S): Abstract 5023.

Castellano D, del M, X, Perez-Gracia JL, Gonzalez-Larriba JL, Abrio MV, et al. (2009) Patient-reported outcomes in a phase III, randomized study of sunitinib versus interferon-{alpha} as first-line systemic therapy for patients with metastatic renal cell carcinoma in a European population. Ann Oncol. 20(11): 1803-1812.

Cella D, Bushmakin AG, Cappelleri JC, Charbonneau C, Li JZ, et al. (2008a) HEALTH-RELATED QUALITY OF LIFE (HRQOL) IN METASTATIC RENAL CELL CARCINOMA (MRCC) PATIENTS RECEIVING SUNITINIB (SU) OR INTERFERON (IFN)-ALFA IN A RANDOMIZED PHASE III TRIAL: UPDATED GEOGRAPHIC ANALYSIS. Ann Oncol. 19(Supplement 8): Abstract 585PD.

Cella D, Li JZ, Bushmakin AG, Cappelleri JC, Kim ST, et al. (2007a) Health-related quality of life (HRQOL) and kidney cancer-related symptoms in patients w ith metastatic renal cell carcinoma (mRCC) treated w ith sunitinib versus interferon (IFN)-alfa: results for
European and US subsample analyses in a randomized, multina. European Journal of Cancer Supplements. 5(4): 114 (Abstract P#1108).

Cella D, Li JZ, Cappelleri JC, Bushmakin A, Charbonneau C, et al. (2008b) Quality of life in patients with metastatic renal cell carcinoma treated with sunitinib or interferon alfa: results from a phase III randomized trial. J Clin Oncol. 26(22): 3763-3769.

Cella D, Michaelson MD, Bushmakin AG, Cappelleri JC, Charbonneau C, et al. (2009) 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 Supplements. 7(2): 175.

Cella D, Yount S, Brucker PS, Du H, Bukowski R, et al. (2007b) Development and validation of a scale to measure disease-related symptoms of kidney cancer. Value Health. 10(4): 285-293.

Cella D, Yount S, Du H, Dhanda R, Gondek K, et al. (2006) Development and validation of the Functional Assessment of Cancer Therapy-Kidney Symptom Index (FKSI). J Support Oncol. 4(4): 191-199.

CHMP. (2005) Guideline on the Evaluation of Anticancer Medicinal Products in Man.

Cockman ME, Masson N, Mole DR, Jaakkola P, Chang GW, et al. (2000) Hypoxia inducible factor-alpha binding and ubiquitylation by the von Hippel-Lindau tumor suppressor protein. J Biol Chem. 275(33): 25733-25741.

Cohen HT, McGovern FJ. (2005) Renal-cell carcinoma. N Engl J Med. 353(23): 2477-2490.

de Souza P., Maart K, Laurell A, Hawkins RE, Berkenblit A, et al. (2007) Results of a phase 3, randomized study of patients w ith advanced renal cell carcinoma (RCC) and poor prognostic features treated with temsirolimus, interferon[alpha] or the combination of temsirolimus + interferon[alpha]. European Journal of Cancer Supplements. 5(4): 282 (Abstract P#4011).

de Souza P, Radulovic S, Beck J, Pendergrass KB, Siddique N, et al. (2008) Characterization of hyperglycemia, hypercholesterolemia, and hyperlipidemia in patients with advanced renal cell carcinoma treated with temsirolimus or interferon-alpha. J Clin Oncol. 26(May 20 Supplement): Abstract 5116.

Dhanda R, Gondek K, Song J, Cella RM, Bukowski RM, et al. (2006) A comparison of quality of life and symptoms in kidney cancer patients receiving sorafenib versus placebo. J Clin Oncol. 24(18S): Abstract 4534.

Drummond MF, Jefferson TO. (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 313(7052): 275-283.

Dutcher J, Szczylik C, Tannir N, Benedetto P, Ruff P, et al. (2007) Correlation of survival with tumor histology, age, and prognostic risk group for previously untreated patients with advanced renal cell carcinoma (adv RCC) receiving temsirolimus (TEMSR) or interferon-alpha (IFN). J Clin Oncol. 25(18S): Abstract 5033.

Dutcher JP, de Souza P., Figlin A, Berkenblit A, Thiele A, et al. Effect of temsirolimus versus interferon-a on survival of patients with advanced renal cellcarcinoma of different tumor histologies. ASCO-GU Abstract 384.

Dutcher JP, de SP, McDermott D, Figlin RA, Berkenblit A, et al. (2009) Effect of temsirolimus versus interferon-alpha on outcome of patients with advanced renal cell carcinoma of different tumor histologies. Med Oncol. 26(2): 202-209.

Eberhardt W, Grunwald V, Ringhoffer.M., Jakse G, Hutson TE, et al. (2006) Phase III randomized trial of sunitinib malate (SU11248) versus interferon-alfa (IFN-alpha) as first-line systemic therapy for patients with metastatic renal cell carcinoma. Onkologie. 29(Supplement 3): 144 (Abstract P650).

Eberhardt W, Grunwalk V, Ringhoffer.M., Brehmer.B., Cella D, et al. (2007) Quality of life predicts for progression-free survival in patients with metastatic renal cell carcinoma treated with sunitinib vs. interferon-alpha. Onkologie. 30(Supplement 3): 17 (Abstract V82).

Egger, M, Davey Smith, G, Altman, D. (2001) Systematic Reviews in Health Care. 2nd Edition ed. London: BMJ Books.

Eisen T, Bukowski R, Staehler M, Szczylik C, Oudard S, et al. (2006) Randomized phase III trial of sorafenib in advanced renal cell carcinoma (RCC): Impact of crossover on survival. J Clin Oncol. 24(18S): Abstract 4524.

Eisen T, Oudard S, Szczylik C, Gravis G, Heinzer H, et al. (2008) Sorafenib for older patients with renal cell carcinoma: subset analysis from a randomized trial. J Natl Cancer Inst. 100(20): 1454-1463.

Escudier B, Bellmunt J, Negrier S, Melichar B, Bracarda S, et al. (2009a) Final results of the phase III, randomized, double blind AVOREN trial of first-line bevacizumab (BEV) + interferon- alpha2a (IFN) in metastatic renal cell carcinoma (mRCC). J Clin Oncol. 27(15S): Abstract 5020.

Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, et al. (2007a) Sorafenib in advanced clear-cell renal-cell carcinoma. N Engl J Med. 356(2): 125-134.

Escudier B, Eisen T, Stadler WM, Szczylik C, Oudard S, et al. (2009b) 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. 27(20): 3312-3318.

Escudier B, Koralewski P, Pluzanska A, Ravaud A, Bracarda S, et al. (2007b) A randomized, controlled, double-blind phase III study (AVOREN) of bevacizumab/interferon-alpha2a vs placebo/interferon- alpha2a as first-line therapy in metastatic renal cell carcinoma. J Clin Oncol. 25(18S): Abstract #3.

Escudier B, Pluzanska A, Koralewski P, Ravaud A, Bracarda S, et al. (2007c) Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. Lancet. 370(9605): 2103-2111.

Escudier B, Ravaud A, Bracarda S, Melichar B, Delva B, et al. Efficacy and safety of first-line bevacizumab (BEV) plus interferona2a (IFN) in subgroups of patients (pts) with metastatic renal cell carcinoma (mRCC). ASCO Abstract 358.

Escudier B, Ravaud A, Negrier S, Szczylik C, Bellmunt J, et al. (2008b) Update on AVOREN trial in metastatic renal cell carcinoma (mRCC): Efficacy and safety in subgroups of patients (pts) and pharmacokinetic (PK) analysis. J Clin Oncol. 26(May 20 Supplement): Abstract 5025.

Escudier B, Szczylik C, Demkow T, Staehler M, Rolland F, et al. (2006) Randomized phase II trial of the multi-kinase inhibitor sorafenib versus interferon (IFN) in treatment-naive patients with metastatic renal cell carcinoma (mRCC). J Clin Oncol. 24(18S): 4501.

Escudier B, Szczylik C, Eisen T, Stadler WM, Schwartz B, et al. (2005) Randomized Phase III trial of the Raf kinase and VEGFR inhibitor sorafenib (BAY 43-9006) in patients with advanced renal cell carcinoma (RCC). J Clin Oncol. 23(16S): Abstract 4510.

Escudier B, Szczylik C, Hutson TE, Demkow T, Staehler M, et al. (2009c) Randomized phase II trial of first-line treatment with sorafenib versus interferon Alfa-2a in patients with metastatic renal cell carcinoma. J Clin Oncol. 27(8): 1280-1289.

Figlin A, Hutson P, Tomczak P, Michaelson MD, Bukowski RM, et al. (2008) Overall survival with sunitinib versus interferon (IFN)-alfa as first-line treatment of metastatic renal cell carcinoma (mRCC). J Clin Oncol. 26(May Supplement): Abstract 5024.

Figlin RA, de SP, McDermott D, Dutcher JP, Berkenblit A, et al. (2009) Analysis of PTEN and HIF-1alpha and correlation with efficacy in patients with advanced renal cell carcinoma treated with temsirolimus versus interferon-alpha. Cancer. 115(16): 3651-3660.

Garcia JA, Rini BI. (2007) Recent progress in the management of advanced renal cell carcinoma. CA Cancer J Clin. 57(2): 112-125.

GlaxoSmithKline. (2008) A Randomized, Double-blind, Placebo-controlled, Multi-center Phase III Study to Evaluate the Efficacy and Safety of Pazopanib (GW786034) Compared to Placebo in Patients with Locally Advanced and/or Metastatic Renal Cell Carcinoma.

Gore ME, Griffin CL, Hancock B, Patel PM, Pyle L, et al. (2010) Interferon alfa-2a versus combination therapy with interferon alfa-2a, interleukin-2, and fluorouracil in patients with untreated metastatic renal cell carcinoma (MRC RE04/EORTC GU 30012): an open-label randomised trial. Lancet. 375(9715): 641-648.

Hancock B, Griffiths G, Ritchie A, Gore M, Mead G, et al. Updated results of the MRC randomised controlled trial of alpha interferon vs MPA in patients with metastatic renal carcinoma. Proceedings of the American Society of Clinical Oncology Abstract 1336.

Harris R, Bradburn M, Deeks J, Harbord R, Altman D. Metan: Stata module for fixed and random effects meta-analysis.
<http://ideas.repec.org/c/boc/bocode/s456798.html>

Harrison ML, Montes A, Gore ME. (2007) New drug therapies for advanced renal cell carcinoma. Expert Rev Anticancer Ther. 7, 57-71.

Hawkins R, Hodge R, Chen M, Neary M, Pickard AS, et al. (2009a) Genitourinary malignancies - Renal cancer Quality of life (QOL) in treatment naïve and cytokine-pretreated patients with advanced renal cell carcinoma (RCC) treated with pazopanib: results from a phase III double-blind, placebo-controlled trial. European Journal of Cancer Supplements. 7(2): 428 (Abstract P-7119).

Hawkins R, Hong SJ, Ulys A, Rolski J, Hong B, et al. (2009b) An open-label extension study to evaluate safety and efficacy of pazopanib in patients with advanced renal cell carcinoma (RCC). J Clin Oncol. 27(15S): Abstract 5110.

Hernberg M, Muhonen T, Pyrhonen S. (1997) Can the CD4+/CD8+ ratio predict the outcome of interferon-alpha therapy for renal cell carcinoma? Ann Oncol. 8(1): 71-77.

Hudes G, Carducci M, Tomczak P, Dutcher J, Figlin R, et al. (2007) Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. N Engl J Med. 356(22): 2271-2281.

Hutson TE, Bellmunt J, Porta C, Staehler M, Szczylik C, et al. (2009a) Long-term safety of sorafenib (SOR) for the treatment (tx) of advanced clear-cell renal-cell carcinoma (RCC): Data analysis from patients (pts) treated for over 1 year in the phase III TARGET study. J Clin Oncol. 27(Supplement): Abstract e16057.

Hutson TE, Bellmunt J, Porto C, Staehler M, Szczylik C, et al. Long-term safety with sorafenib in advanced RCC patients: Data analysis from patients treated for over 1 year in TARGET trial for the Sorafenib TARGET Clinical Trial Group. ASCO-GU Abstract 326.

Hutson TE, Davis ID, Machiels JP, De Souza PL, Hong BF, et al. (2007) Pazopanib (GW786034) is active in metastatic renal cell carcinoma (RCC): Interim results of a phase II randomized discontinuation trial (RDT). J Clin Oncol. 25(18S): Abstract 5031.

Hutson TE, Davis ID, Machiels JP, De Souza PL, Rottey S, et al. (2010) Efficacy and safety of pazopanib in patients with metastatic renal cell carcinoma. J Clin Oncol. 28(3): 475-480.

Iliopoulos O, Levy AP, Jiang C, Kaelin WG, Jr., Goldberg MA. (1996) Negative regulation of hypoxia-inducible genes by the von Hippel-Lindau protein. Proc Natl Acad Sci U S A. 93(20): 10595-10599.

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, et al. (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 17(1): 1-12.

Jager E, Heinzer H, Grimm M-O, Krause S, Scheuring U, et al. (2005) Randomized phase III trial of the multiple kinase inhibitor sorafenib (BAY 43-9006) in patients with advanced renal cell carcinoma (RCC). Onkologie. 28(Supplement 3): 1.

Kriegmair M, Oberneder R, Hofstetter A. (1995) Interferon alfa and vinblastine versus medroxyprogesterone acetate in the treatment of metastatic renal cell carcinoma. Urology. 45(5): 758-762.

Larkin MG, Kipps ELS, Powell CJ, Swanton C. (2009) Review: Systemic therapy for advanced renal cell carcinoma. Therapeutic Advances in Medical Oncology. 1(1): 15-27.

Lasset C, Savary J, Negrier S. (1992) Early severe toxicity of immunotherapy combining IL2 and IFN in patients with metastatic renal cell carcinoma included in the French CRECY study. Ann Oncol. 3(Supplement 5): 140 (Abstract 541).

Lee EH, Chun M, Kang S, Lee HJ. (2004) Validation of the Functional Assessment of Cancer Therapy-General (FACT-G) scale for measuring the health-related quality of life in Korean women with breast cancer. Jpn J Clin Oncol. 34(7): 393-399.

Logan T, McDermott DF, Dutcher JP, Makhson A, Berkenblit A, et al. (2008) Exploratory analysis of the influence of nephrectomy status on temsirolimus efficacy in patients with advanced renal cell carcinoma and poor-risk features. J Clin Oncol. 26(May 20 Supplement): Abstract 5050.

Mallick R, Chen J. (2008) PREDICTORS OF SURVIVAL IN PATIENTS WITH ADVANCED RENAL CELL CARCINOMA WHO RECEIVED FIRST LINE TREATMENT WITH TEMSIROLIMUS, INTERFERON-A OR COMBINATION TEMSIROLIMUS/INTERFERON-A. Ann Oncol. 19(Supplement 8): Abstract 591P.

Melichar B, Bracarda S, Bellmunt J, Ravaud A, Negrier S, et al. (2009) First-line bevacizumab + reduced dose interferon-alpha2a in patients (pts) with metastatic renal cell carcinoma (mRCC): an update on overall survival. European Journal of Cancer Supplements. 7(2): 430 (Abstract P-7123).

Melichar B, Koralewski P, Pluzanska A, Ravaud A, Bracarda S, et al. (2007) First-line bevacizumab improves progression-free survival with lower doses of interferon[alpha]2a in the treatment of patients with metastatic renal cell carcinoma (AVOREN). European Journal of Cancer Supplements. 5(4): 304 (Abstract P#4518).

Melichar B, Koralewski P, Ravaud A, Pluzanska A, Bracarda S, et al. (2008) First-line bevacizumab combined with reduced dose interferon-alpha2a is active in patients with metastatic renal cell carcinoma. Ann Oncol. 19(8): 1470-1476.

Mickisch G, Gore M, Escudier B, Procopio G, Walzer S, et al. (2009) 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., 1-7.

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, et al. (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 354(9193): 1896-1900.

Moher D, Liberati A, Tetzlaff J, Altman DG. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6(7): e1000097.

Moore L. (2006) A phase 3, randomized, 3-arm study of temsirolimus (TEMSR) or interferon-alpha (IFN) or the combination of TEMSR + IFN in the treatment of first-line, poor-risk patients with advanced renal cell carcinoma (adv RCC). J Clin Oncol. 24(18S): LBA4.

Motzer RJ. (2006a) Sunitinib superior to interferon-(alpha) in metastatic kidney cancer. P&T. 31(9): 539-540.

Motzer RJ, Bukowski RM, Figlin RA, Hutson TE, Michaelson MD, et al. (2008) Prognostic nomogram for sunitinib in patients with metastatic renal cell carcinoma. Cancer. 113(7): 1552-1558.

Motzer RJ, Figlin RA, Hutson TE, Tomczak P, Bukowski RM, et al. Sunitinib versus interferon-alfa (IFN-a) as first-line treatment of metastatic renal cell carcinoma (mRCC): Updated results and analysis of prognostic factors. ASCO-GU Abstract 5024.

Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, et al. (2009) Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. J Clin Oncol. 27(22): 3584-3590.

Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, et al. (2006b) Phase III randomized trial of sunitinib malate (SU11248) versus interferon-alfa (IFN-{alpha}) as first-line systemic therapy for patients with metastatic renal cell carcinoma (mRCC). J Clin Oncol. 24(18S): Abstract LBA3.

Motzer RJ, Hutson TE, Tomczak P, Michaelson MD, Bukowski RM, et al. (2007b) Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. N Engl J Med. 356(2): 115-124.

Motzer RJ, Mazumdar M, Bacik J, Berg W, Amsterdam A, et al. (1999) Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma. J Clin Oncol. 17(8): 2530-2540.

Motzer RJ, Michaelson MD, Hutson TE, Tomczak P, Bukowski RM, et al. (2007c) Sunitinib versus interferon (IFN)-alfa as first-line treatment of metastatic renal cell carcinoma (mRCC): updated efficacy and safety results and further analysis of prognostic factors. European Journal of Cancer Supplements. 5(4): 301 (Abstract P#4509).

Negrier S, Escudier B, Lasset C, Douillard JY, Savary J, et al. (1998) Recombinant human interleukin-2, recombinant human interferon alfa-2a, or both in metastatic renal-cell carcinoma. Groupe Francais d'Immunotherapie. N Engl J Med. 338(18): 1272-1278.

Negrier S, Escudier B, Lasset C, Savary J, Douillard JY, et al. The FNCLCC Crecy trial: interleukin 2 (IL2) + interferon (IFN) is the optimal treatment to induce responses in metastatic renal cell carcinoma (MRCC). Proceedings of the American Society of Clinical Oncology Abstract 629.

Negrier S, Figlin RA, Hutson TE, Tomczak P, Michaelson MD, et al. (2008) OVERALL SURVIVAL WITH SUNITINIB VERSUS INTERFERON (IFN)-ALFA AS FIRST-LINE TREATMENT OF METASTATIC RENAL CELL CARCINOMA (MRCC). Ann Oncol. 19(8): Abstract 588PD.

Negrier S, Jager E, Porta C, McDermott D, Moore M, et al. (2009) Efficacy and safety of sorafenib in patients with advanced renal cell carcinoma with and without prior cytokine therapy, a subanalysis of TARGET. Med Oncol.

Negrier S, Perol D, Ravaud A, Chevreau C, Bay JO, et al. (2007) Medroxyprogesterone, interferon alfa-2a, interleukin 2, or combination of both cytokines in patients with metastatic renal carcinoma of intermediate prognosis: results of a randomized controlled trial. Cancer. 110(11): 2468-2477.

NICE. TA 169 Sunitinib for the first-line treatment of advanced and/or metastatic renal cell carcinoma. <http://guidance.nice.org.uk/TA169> Accessed Dec. 2009a.

NICE. TA 178 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 . <http://www.nice.org.uk/TA178> Accessed Dec. 2009b.

Oudard S, Eisen T, Szczylik C, Siebels MS, Negrier S, et al. (2009) 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. 27(Supplement): Abstract e16099.

Parasuraman S, Hudes G, Levy D, Strahs A, Moore L, et al. (2007) Comparison of quality-adjusted survival in patients with advanced renal cell carcinoma receiving first-line treatment with temsirolimus (TEMSR) or interferon-alpha (IFN) or the combination of IFN+TEMSR. J Clin Oncol. 25(18S): Abstract 5049.

Park JO, Lee SI, Song SY, Kim K, Kim WS, et al. (2003) Measuring response in solid tumors: comparison of RECIST and WHO response criteria. Jpn J Clin Oncol. 33(10): 533-537.

Pascual D, Borque A. (2008) Epidemiology of kidney cancer. Adv Urol., 782381.

Patil S, Hutson TE, Figlin A, Hutson, Michaelson MD, et al. (2009) Prognostic factors for overall survival with sunitinib as first-line therapy in patients with metastatic renal cell carcinoma (mRCC). J Clin Oncol. 27(15S): Abstract 5042.

Pendergrass KB, Hudes G, Radulovic S, Beck J, Dutcher JP, et al. Characterization of hyperglycemia, hypercholesterolemia and hyperlipidemia in patients with advanced renal cell carcinoma treated with temsirolimus or interferon -a. ASCO-GU Abstract 297.

Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. (2006) Comparison of two methods to detect publication bias in meta-analysis. JAMA. 295(6): 676-680.

Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, et al. (2004) Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess. 8(36): iii-xi, 1.

Prummer O, Porzsolt F. (1994) Recombinant interferon-alpha 2 antibodies in renal cell carcinoma. Delta-P Study Group. J Interferon Res. 14(4): 193-195.

Pyrhonen S, Salminen E, Ruutu M, Lehtonen T, Nurmi M, et al. (1999) Prospective randomized trial of interferon alfa-2a plus vinblastine versus vinblastine alone in patients with advanced renal cell cancer. J Clin Oncol. 17(9): 2859-2867.

Rajagopalan S, Pullenayegum E, Alemao E, Strahs A, Purvis J. (2009) Evaluation of adverse event –related hospitalizations in patients with advanced renal cell carcinoma on treatment with temsirolimus or interferon-alfa: results from a phase 3 randomized trial. European Journal of Cancer Supplements. 7(2): 441 (Abstract P-7160).

Ravaud A, Bajetta E, Bracarda S, Negrier S, Szczylik C, et al. (2008) Long-term tolerability of first-line bevacizumab (BEV) administered for > 12 months to patients (PTS) with metastatic renal cell carcinoma (MRCC). Ann Oncol. 19(Supplement 8): Abstract 5790.

Reddy K. (2006) Phase III study of sunitinib malate (SU11248) versus interferon-alpha as first-line treatment in patients with metastatic renal cell carcinoma. Clin Genitourin Cancer. 5(1): 23-25.

Remak E, Charbonneau C, Negrier S, Kim ST, Motzer RJ. (2008) Economic evaluation of sunitinib malate for the first-line treatment of metastatic renal cell carcinoma. J Clin Oncol. 26(24): 3995-4000.

Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, et al. (2009) Bevacizumab plus interferon-alpha versus interferon-alpha monotherapy in patients with metastatic renal cell carcinoma: Results of overall survival for CALGB 90206. J Clin Oncol. 27(18S): Abstract LBA5019.

Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, et al. (2008a) Bevacizumab plus interferon alfa compared with interferon alfa monotherapy in patients with metastatic renal cell carcinoma: CALGB 90206. J Clin Oncol. 26(33): 5422-5428.

Rini BI, Halabi S, Rosenberg JE, Stadler WM, Vaena DA, et al. CALGB 90206: A phase III trial of bevacizumab plus interferon-alpha versus interferon-alpha monotherapy in metastatic renal cell carcinoma. ASCO-GU Abstract 350.

Rini BI, Halabi S, Taylor J, Small EJ, Schilsky RL. (2004) Cancer and Leukemia Group B 90206: A randomized phase III trial of interferon-alpha or interferon-alpha plus anti-vascular endothelial growth factor antibody (bevacizumab) in metastatic renal cell carcinoma. Clin Cancer Res. 10(8): 2584-2586.

Ritchie A, Griffiths G, Cook P, Oliver RT, Hancock B, et al. Alpha interferon improves survival in patients with metastatic renal carcinoma-preliminary results of an MRC randomised controlled trial. Proceedings of the American Society of Clinical Oncology Abstract 1194.

Ritchie AWW, Griffiths G, Parmar M. (1999) Interferon-alpha and survival in metastatic renal carcinoma: early results of a randomised controlled trial. Medical Research Council Renal Cancer Collaborators. Lancet. 353(9146): 14-17.

Royston P, Parmar MK, Altman DG. (2008) Visualizing length of survival in time-to-event studies: a complement to Kaplan-Meier plots. J Natl Cancer Inst. 100(2): 92-97.

Royston P, Sauerbrei W, Ritchie A. (2004) Is treatment with interferon-alpha effective in all patients with metastatic renal carcinoma? A new approach to the investigation of interactions. Br J Cancer. 90(4): 794-799.

Sharp SJ, Thompson SG, Altman DG. (1996) The relation between treatment benefit and underlying risk in meta-analysis. BMJ. 313(7059): 735-738.

Soret JY, Escudier B. (1996) Adjuvant treatment with IL-2 or interferon-alpha in renal cell carcinoma: a French multicentric study. Cancer Biother Radiopharm. 11(5): 301-302.

Steineck G, Strander H, Carbin BE, Borgstrom E, Wallin L, et al. (1990) Recombinant leukocyte interferon alpha-2a and medroxyprogesterone in advanced renal cell carcinoma. A randomized trial. Acta Oncol. 29(2): 155-162.

Sternberg CN, Davis I, Wagstaff J, Hawkins R, Chen M, et al. (2009a) Predictive and prognostic factors in a phase III study of pazopanib in patients with advanced renal cell carcinoma (RCC). European Journal of Cancer Supplements. 7(2): 424 (Abstract O-7106).

Sternberg CN, Szczylik C, Lee E, Salman PV, Mardiak J, et al. (2009b) A randomized, double-blind phase III study of pazopanib in treatment-naive and cytokinepretreated patients with advanced renal cell carcinoma (RCC). J Clin Oncol. 27(15S): Abstract 5021.

Szczylik C, Demkow T, Staehler M, Rolland F, Negrier S, et al. (2007) Randomized phase II trial of first-line treatment with soratinib versus interferon in patients with advanced renal cell carcinoma: Final results. J Clin Oncol. 25(18S): 5025.

Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, et al. (2000) New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 92(3): 205-216.

Thompson SG. (1994) Why sources of heterogeneity in meta-analysis should be investigated. BMJ. 309(6965): 1351-1355.

Williamson PR, Gamble C, Altman DG, Hutton JL. (2005) Outcome selection bias in meta-analysis. Stat Methods Med Res. 14(5): 515-524.

Yang S, Hudes G, de Souza P, Alemao E, Strahs A, et al. (2009) Evaluation of quality of life in patients with advanced renal cell carcinoma treated with temsirolimus vs interferon-alfa: results from a phase III randomized trial. European Journal of Cancer Supplements. 7(2): 433 (Abstract P-7134).