Bevacizumab, sorafenib tosylate, sunitinib and temsirolimus for renal cell carcinoma: a systematic review and economic evaluation

J Thompson Coon,[1]* M Hoyle,[1] C Green,[1] Z Liu,[1] K Welch,[2] T Moxham[1] and K Stein[1]

[1]Peninsula Technology Assessment Group (PenTAG), Peninsula College of Medicine and Dentistry, University of Exeter, UK
[2]Wessex Institute for Health Research and Development (WIHRD), School of Medicine, University of Southampton, UK

*Corresponding author

Executive summary

Health Technology Assessment 2010; Vol. 14: No. 2
DOI: 10.3310/hta14020

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk

Executive summary

Background

Renal cell carcinoma (RCC) is a highly vascular type of kidney cancer arising in the epithelial elements of the nephrons. The most common histological subtype of RCC is clear cell carcinoma (approximately 75% of cases). RCC is often asymptomatic until it reaches a late stage. In England and Wales, kidney cancer is the eighth most common cancer in men and the fourteenth most common in women. Of those diagnosed with RCC in England and Wales, about 44% live for at least 5 years after initial diagnosis and about 40% for at least 10 years. However, prognosis following diagnosis of metastatic disease is poor, and only about 10% of people diagnosed with stage IV RCC live for at least 5 years after diagnosis.

Current NHS treatment options for metastatic RCC include radical nephrectomy and interferon (IFN). There is currently no standard NHS treatment for patients with metastatic RCC who do not respond to first-line immunotherapy or who are unsuitable for treatment with IFN. Recently developed therapeutic agents include: *bevacizumab*, licensed for use as first-line therapy in patients with advanced and/or metastatic RCC; *sorafenib tosylate*, licensed for first-line therapy in individuals who are not suitable for treatment with IFN and as second-line therapy in those in whom treatment with cytokine-based immunotherapy has failed; *sunitinib*, licensed for use in the first- and second-line treatment of advanced and/or metastatic RCC; and *temsirolimus*, licensed for first-line treatment of patients with advanced RCC who have at least three of six poor prognostic risk factors.

Objectives

To assess the clinical effectiveness and cost-effectiveness of bevacizumab combined with IFN, sorafenib tosylate, sunitinib and temsirolimus in the treatment of people with advanced and/or metastatic RCC, specifically:

- to identify, appraise and synthesise the current evidence for the above in accordance with their marketing authorisations
- to determine what, if any, is the incremental cost-effectiveness of the interventions in comparison with current standard treatment.

The report addresses the following policy questions:

1. In those suitable for first-line treatment with immunotherapy: bevacizumab plus IFN versus IFN alone and sunitinib versus IFN alone, using IFN as a comparator.
2. In those not suitable for first-line treatment with immunotherapy: sorafenib and sunitinib, using best supportive care as a comparator.
3. In those with three or more of six poor prognostic factors: bevacizumab plus IFN, sorafenib, sunitinib, temsirolimus and best supportive care, using IFN as a comparator.
4. In those in whom cytokine based immunotherapy has failed: second-line therapy with sorafenib and sunitinib, using best supportive care as a comparator.

Methods

Clinical effectiveness systematic review

Electronic databases, including MEDLINE, EMBASE and the Cochrane Library, were searched up to September/October 2007 (and again in February 2008). Systematic reviews and randomised clinical trials comparing any of the interventions with any of the comparators in participants with advanced and/or metastatic RCC were included. The use of data from phase II studies and non-randomised clinical trials was considered where there was insufficient evidence from good-quality randomised clinical trials. Conference abstracts were included if there was sufficient detail to adequately assess quality. Full papers for studies that appeared relevant were retrieved and screened in detail. All trials were fully data extracted and quality assessed. Results of the included trials were synthesised narratively. The validity of indirect comparison between interventions was considered, using the method proposed by Bucher and colleagues, where data from head-to-head randomised clinical trials were unavailable.

Review of economic evaluations, related literature and manufacturer submissions

Electronic databases were searched up to September/October 2007 (and again in March 2008). All titles and abstracts were assessed independently and all publications meeting the inclusion criteria were fully data extracted and discussed narratively. Searches were also performed to identify literature describing health-related quality of life of people with RCC, treatment costs and resource use associated with the treatment of RCC, and modelling methods used to model disease progression and cost-effectiveness in RCC. The cost-effectiveness analyses reported in the manufacturers' submissions were assessed against the NICE reference case and critically

appraised using the framework presented by Phillips and colleagues.

PenTAG cost–utility model

A decision-analytic Markov-type model was developed in EXCEL to simulate disease progression and estimate the cost-effectiveness of the drugs under consideration. The model has three health states – progression-free survival, progressive disease, and death – and uses estimates of effectiveness, costs and health-state utilities assigned to these states to model disease progression and cost-effectiveness over time. Future costs and benefits were discounted at 3.5% per annum. Weibull survival curves were fitted to the progression-free and overall survival Kaplan–Meier curves from clinical trials for the baseline comparator. Relative measures of treatment effectiveness (hazard ratios, HRs) were then used to estimate the expected disease progression compared with baseline. One-way, multi-way and probabilistic sensitivity analyses were used to explore structural and parameter uncertainty.

Results

Number and quality of effectiveness studies

A total of 888 titles and abstracts were retrieved. Thirteen publications describing eight clinical trials were included. Of these, seven were fully published randomised clinical trials and one was a protocol and conference abstract. Data contained within a further 19 conference abstracts relating to the included trials were also considered.

Three randomised clinical trials were identified that compared either bevacizumab plus IFN (two trials, one published in abstract form only) or sunitinib (one trial) with IFN alone as first-line therapy in those suitable for treatment with IFN. Preliminary results (abstract only) of one randomised clinical trial in which sorafenib tosylate was compared with best supportive care in people unsuitable for treatment with IFN, and one randomised clinical trial of temsirolimus versus IFN in people with three or more of six risk factors for poor prognosis were located. For second-line therapy, we found a randomised clinical trial and a randomised discontinuation trial of sorafenib versus best supportive care and two phase II single-arm trials of sunitinib.

We were unable to identify any data on clinical effectiveness in the following areas: sunitinib or best supportive care in patients unsuitable for treatment with immunotherapy; sorafenib in patients with poor prognosis; or sunitinib as second-line therapy.

All the fully published included studies were large, multicentre, good-quality trials. There was insufficient detail in the conference abstracts to fully appraise the quality of the trials.

Summary of benefits and risks

Bevacizumab plus IFN and sunitinib compared with IFN as first-line therapy Treatment with both interventions had clinically relevant and statistically significant advantages over treatment with IFN alone, in terms of progression-free survival and tumour response, doubling median progression-free survival from approximately 5 months to 10 months. There was insufficient data on overall survival due to the early crossover of patients on control treatment following interim analyses; however, both interventions showed some benefits in terms of overall survival. An indirect comparison between sunitinib and bevacizumab plus IFN suggested that sunitinib may be more effective than bevacizumab plus IFN [HR 0.67; 95% confidence interval (CI) 0.50 to 0.89) in terms of progression-free survival. Sunitinib was associated with a lower frequency of adverse events than IFN, and bevacizumab plus IFN with slightly more than IFN alone.

Sorafenib tosylate and sunitinib compared with best supportive care as first-line therapy No trials met the inclusion criteria.

Bevacizumab plus IFN, sorafenib, sunitinib, temsirolimus and best supportive care compared with IFN as first-line therapy in people with poor prognosis Temsirolimus had clinically relevant and statistically significant advantages over treatment with IFN in terms of progression-free and overall survival, increasing median overall survival from 7.3 to 10.9 months (HR 0.73; 95% CI 0.58 to 0.92). There was also evidence to suggest that progression-free survival may be prolonged by treatment with the combination of bevacizumab plus IFN compared with IFN alone, though it is not clear whether this effect would be considered clinically and statistically significant. We were unable to find any data on sorafenib in this population. A significantly lower frequency of grade 3 and 4 adverse events was reported with temsirolimus than with IFN.

Sorafenib tosylate and sunitinib compared with best supportive care as second-line therapy Sorafenib had clinically relevant and statistically significant advantages over best supportive care in terms of overall survival, progression-free survival and tumour response, with progression-free survival doubling in the randomised clinical trial (HR 0.51; 95% CI 0.43 to 0.60). However, it was associated with an increased frequency of hypertension and hand–foot skin reaction compared with placebo. We were unable to locate any comparative trials of sunitinib as second-line therapy, but two single-arm phase II trials suggested that sunitinib may be efficacious in this population.

Summary of cost-effectiveness

We were unable to locate any fully published economic evaluations of any of the interventions. Although there are many similarities in the methodology and structural

assumptions employed by Peninsula Technology Assessment Group (PenTAG) and the manufacturers of the interventions, in all cases the cost-effectiveness estimates from the PenTAG economic evaluation were higher than those presented in the manufacturers' submissions.

Bevacizumab plus IFN and sunitinib compared with IFN as first-line therapy The PenTAG model estimated that the cost per quality-adjusted life-year (QALY) for bevacizumab plus IFN versus IFN is £171,301. If the NHS is willing to pay £30,000 for an additional QALY, there is zero probability that this intervention would be considered cost-effective, and bevacizumab plus IFN is unlikely to be considered cost-effective at any reasonable willingness-to-pay threshold. For sunitinib versus IFN, the PenTAG model estimated a cost per QALY of £71,462. Sunitinib is likely to be considered cost-effective compared with both bevacizumab plus IFN and IFN alone only above a willingness-to-pay threshold of £75,000 per QALY.

Sorafenib tosylate and sunitinib compared with best supportive care as first-line therapy Insufficient clinical effectiveness data to perform a cost-effectiveness analysis.

Bevacizumab plus IFN, sorafenib, sunitinib, temsirolimus and best supportive care compared with IFN as first-line therapy in people with poor prognosis We were unable to locate appropriate overall and progression-free survival data with which to populate an economic model for the first three interventions or best supportive care. The base-case discounted incremental cost-effectiveness ratio (ICER) for temsirolimus versus IFN estimated from the PenTAG model was £81,687 per QALY. Temsirolimus is likely to be considered cost-effective compared with IFN only above a willingness-to-pay threshold of £82,000 per QALY. The cost–utility analyses performed in patient subgroups indicate cost per QALY estimates ranging from £64,680 to £132,778, although the clinical effectiveness data on which these analyses are based is uncertain.

Sorafenib tosylate and sunitinib compared with best supportive care as second-line therapy We were unable to locate any comparative trials of sunitinib as second-line therapy in this population. The PenTAG model estimated a cost per QALY for sorafenib versus best supportive care of £102,498. Compared with best supportive care, sorafenib is only likely to be considered cost-effective above a willingness-to-pay threshold of approximately £100,000 per QALY.

Sensitivity analyses

In all comparisons, the cost-effectiveness estimates were particularly sensitive to variations in the estimates of treatment effectiveness, drug pricing (including dose intensity data), and health-state utility input parameters. The ICERs were insensitive to a number of assumptions and data estimates, in particular discounting, time horizon, limiting IFN administration to 1 year, non-drug costs, inclusion of estimates associated with costs of death, and estimates of adverse event costs.

Discussion

The assessment was necessarily constrained by the marketing authorisations of the interventions under review, leading to difficulties in deriving research questions applicable to the population with RCC. We felt that it was important to use current standard treatment as the comparator wherever possible – considering IFN to be the comparator for first-line therapy in patients suitable for treatment with immunotherapy and best supportive care the comparator in all other situations. Suitability for treatment with immunotherapy was defined in terms of clinical contraindication to treatment (e.g. autoimmune disease or a history of depression). However, we acknowledge that a large proportion of people diagnosed with RCC in the UK will be deemed unsuitable for treatment with IFN as a result of clinical markers of prognosis. Informal extrapolation of available data suggests that if it is assumed that there is no difference in the relative effectiveness of best supportive care and IFN in this population, and that the cost of best supportive care would be less than the cost of treatment with IFN, it is possible that the new interventions would be less likely to be considered cost-effective at commonly used willingness-to-pay thresholds when compared with best supportive care.

Clinical trials suggested that all four interventions have clinically relevant and statistically significant advantages over current standard treatment (IFN or best supportive care) where data exists with which to make the comparison. The most robust clinical effectiveness data was for progression-free survival; treatment crossover following interim analyses was permitted in all but one (temsirolimus versus IFN) of the included trials resulting in confounding of overall survival data. There is therefore a large amount of uncertainty in the estimates used in the assessment of clinical effectiveness and cost-effectiveness.

The PenTAG model estimated that if the NHS is willing to pay £30,000 for an additional QALY, the probability that any of the interventions (in the undertaken comparisons) would be considered cost-effective is zero. Exploration of these results using one-way, multi-way and probabilistic sensitivity analyses indicated that the model is most sensitive to variations in the HRs for overall survival, drug pricing (including assumptions made about dose intensities and drug wastage) and health-state utility values. The sensitivity analyses for the HRs for progression-free survival have highlighted issues linked to the balancing of incremental costs and effects. In the PenTAG analysis, improvements in progression-free survival made the drugs less attractive in terms of value for money. This counterintuitive effect was seen across all of the analyses undertaken by PenTAG, was apparent for both cost per QALY and cost per life-year analyses and could be explained partly by the

relatively high incremental treatment costs (costs of the drug, drug administration and monitoring) associated with time spent in the progression-free disease health state. The cost-effectiveness estimates produced in the PenTAG economic evaluation were higher than the manufacturers' base-case estimates in all cases (although in two of the four analyses the results are similar). Although the manufacturers and PenTAG's analyses share some common aspects of methodology, there are also clear differences in the resulting cost-effectiveness estimates.

Strengths and limitations of the analyses

Strengths include comprehensive, explicit and systematic literature searches, including hand searching of conference proceedings, to locate evidence for the review of clinical effectiveness and inform the economic modelling study; work to fit the most appropriate survival curves to the empirical immature overall survival data; and extensive analyses of the uncertainty of the model using one-way, multi-way and probabilistic sensitivity analyses.

Limitations include the constraints on the assessment by the marketing authorisations of the interventions; the uncertainty of the overall survival and health-state utility data; the availability of clinical effectiveness data for all potential comparisons; issues around patient preference; consideration of the sequencing of treatments; some of the structural modelling assumptions used in the PenTAG model; and the scarcity of available information on resource use and costs.

Generalisability of the findings

All the trials included in the review of clinical effectiveness were conducted in patients with predominantly clear cell, metastatic RCC, the majority of whom had undergone previous nephrectomy and many of whom had favourable and intermediate prognosis and good performance status. None of the studies recruited patients with brain metastases (unless neurologically stable) and few patients with bone metastases were included (20% in the trial of bevacizumab plus IFN versus IFN and 30% in the trial of sunitinib versus IFN). Whether the results of this assessment can be extrapolated to other patient groups is unclear.

Conclusions

Evidence suggests that treatment with bevacizumab plus IFN and sunitinib has clinically relevant and statistically significant advantages over treatment with IFN alone in patients with metastatic RCC. Also, in people with three of six risk factors for poor prognosis, temsirolimus has clinically relevant advantages over treatment with IFN, and sorafenib tosylate is superior to best supportive care as second-line therapy. The frequency of adverse events associated with bevacizumab plus IFN, sunitinib and temsirolimus is comparable with that seen with IFN, although the adverse event profile is different. Treatment with sorafenib is associated with a significantly increased frequency of hypertension and hand–foot syndrome. The PenTAG cost-effectiveness analyses suggest that the probability that any of the interventions would be considered cost-effective at a willingness-to-pay threshold of £30,000 per QALY is zero.

Suggested future research questions and priorities

There are clear gaps in the evidence base needed to fully appraise the clinical effectiveness and cost-effectiveness of these four interventions in accordance with their marketing authorisations:

1. More randomised clinical trials in the following areas would be useful: in patients unsuitable for treatment with IFN because of contraindications or who have been defined as having intermediate and poor prognosis and therefore unlikely to benefit from IFN; studies of sorafenib tosylate, sunitinib, bevacizumab plus IFN and best supportive care; and comparative trials of sunitinib and sorafenib as second-line therapy.
2. Research to improve understanding of the impact of the interventions on health-related quality of life during progression-free survival and progressed disease would facilitate the decision-making process for clinicians and patients.
3. Research on current treatment pathways and practice (e.g. in the use of IFN) would reduce the level of uncertainty in future studies modelling the cost-effectiveness of drugs for treatment of renal cancer.
4. As more treatments are introduced, the issues of treatment sequencing become more important: more research is needed on the combination and order of treatments to provide maximum benefit in each patient population.
5. Modelling treatment of RCC presents methodological challenges when using summary data (survival analysis) from clinical trials: research on the impact of using aggregated versus individual patient-level data would be useful.

Publication

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

NIHR Health Technology Assessment programme

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

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

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

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 07/72/01. The protocol was agreed in November 2007. The assessment report began editorial review in May 2008 and was accepted for publication in March 2009. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

ISSN 1366-5278

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.

Printed on acid-free paper in the UK by the Charlesworth Group.