NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Renal cell carcinoma pathway model report

This document is not NICE guidance. It is a summary of the external assessment group's model and assumptions on renal cell carcinoma, discussed by NICE's technology appraisal committee B. This report is the basis of NICE's final draft guidance on cabozatinib with nivolumab for untreated advanced renal cell carcinoma.

1 Model approach, key assumptions and committee discussion

The evaluation committee used NICE's pilot pathway model approach. This is a single economic model developed by an external assessment group (EAG) evaluating a whole disease area (in this case, renal cell carcinoma [RCC]) and how treatments fit into this. This report is a summary of the committee's decisions about the model starting from first-line treatment of advanced RCC, and is the basis for NICE's final draft guidance on cabozantinib with nivolumab for untreated RCC. This model does not cover early-stage or adjuvant treatment for RCC. Email NICE for access to the model and see the committee papers for the EAG's assessment report.

Committee's preferred assumptions for key issues

1.1 The committee's preferred assumptions were to:

- use a state transition model approach (see section 1.6)
- consider 4 lines of treatment then best supportive care (see section 1.7)
- use a fractional polynomial network meta-analysis to inform treatment efficacy in the model (see section 1.16)
- estimate time to stopping treatment and time to next treatment for comparators by applying hazard ratios from the progression-free survival network meta-analyses to the baseline real-world evidence curves for those parameters (see section 1.21)
- assume equal effectiveness for cabozantinib and sunitinib for first-line progression-free survival for intermediate- and poor-risk cancer
- use the adverse events network meta-analyses to model adverse events for comparators (see section 1.25)
- use the EAG's approach for estimating utility from previously accepted NICE technology appraisals (see section 1.27)
- use relative dose intensities from published clinical trials (see section 1.29).

See table 1 for a summary of all the committee's preferred assumptions.

Pathway model

The condition

Effect on quality of life

1.2 Patient experts explained that advanced RCC is life changing. They explained how RCC affects people's lives, starting from the shock and despair of initial diagnosis. It is difficult for people with RCC to continue with daily life even after successful treatment, because of the fear of disease recurrence. Patient experts said that people with advanced and metastatic RCC are frequently hospitalised, may have to take early retirement and have uncertainty about the future. Commonly there is a substantial psychological impact. Patient experts explained that current treatment options are associated with toxicity, which can result in needing to take time off work. There is inconsistency in which treatment options are available across the country, and for some people there are no treatment options at all. The committee concluded that advanced RCC has a large impact on quality of life.

Population and subgroups

1.3 The committee considered whether the model population and subgroups were appropriate. As per NICE's scope, only advanced RCC (stage 3 unresectable RCC, or stage 4 RCC) was included in the decision problem. Clinical trials and treatment decisions are often guided by risk status. RCC is usually grouped into 2 categories: favourable-risk or intermediate- and poor-risk disease, as defined by the International Metastatic Renal Cell Carcinoma Database Consortium (IMDC) criteria. Clinical experts explained that approximately 80% of people with RCC in the UK have intermediate- or poor-risk cancer, and that this distribution is also seen globally. The committee noted that the IMDC criteria use certain continuous prognostic factors (for example, Karnofsky performance status and time from diagnosis to treatment). This means people may move

between risk groups over time. The committee also explained that clinical trials for RCC have also shown differences in treatment effect between these risk groups. So, to reflect clinical practice and ensure robust decision making, the committee agreed that the 2 subgroups (favourable-risk and intermediate- and poor-risk) should be included in the model, along with the all-risk group (which includes both subgroups). The committee concluded that investigating subgroups by risk status was appropriate for the model.

Treatment pathway

First line

1.4 The committee considered how the EAG had modelled first-line treatments (see figure 1). Currently available first-line treatments for advanced RCC include:

- tyrosine kinase inhibitors (TKIs), including sunitinib, pazopanib, tivozanib and cabozantinib
- immunotherapies with other immunotherapies, including nivolumab plus ipilimumab
- immunotherapies with TKIs, including lenvatinib plus pembrolizumab, or avelumab plus axitinib (available through the Cancer Drugs Fund [CDF]).

Treatment is decided based on risk status (see section 1.3). Nivolumab plus ipilimumab, lenvatinib plus pembrolizumab, and cabozantinib are only available for intermediate- and poor-risk RCC. Clinical experts described what affects treatment decisions. First, healthcare professionals decide whether to treat or not. If treatment is considered, they assess these criteria:

- safety
- comorbidities or whether someone is immunocompromised
- potential side effects

- the expected chance of response, and potential duration.

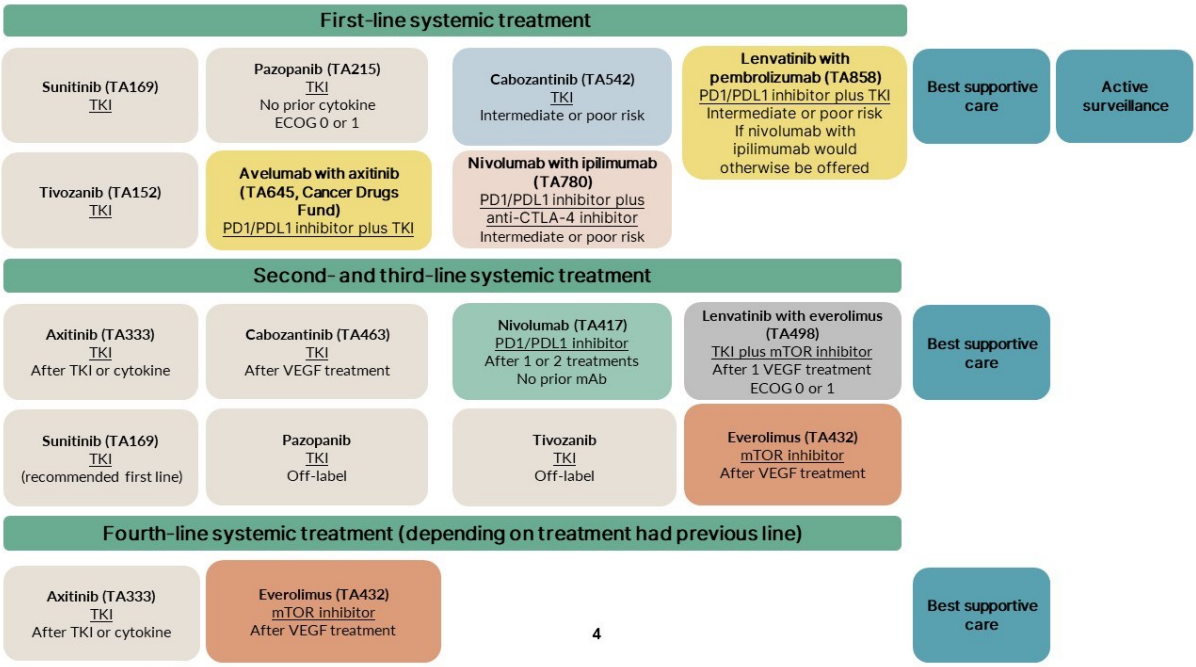
The clinical expert noted that TKI monotherapy use was declining and immunotherapy combination treatment use (which clinicians generally consider to be more effective) was growing for people with intermediate- or poor-risk disease. They noted that tivozanib is the least used of the TKIs. The committee considered that because tivozanib is a NICE-recommended treatment and used in practice, it is relevant to the decision problem. The committee noted that avelumab plus axitinib was only available through the Cancer Drugs Fund so is not considered standard care and should not be included in the current decision problem. The committee concluded that the first-line treatments used in the model represented those used in NHS practice. See section 2.3.2 of the EAG's assessment report for more details on how first-line treatment was modelled.

Later lines

1.5 The committee considered how the EAG had modelled later-line treatments. Second- and third-line treatments are similar to those available at first line (with some used off label) but also include nivolumab, axitinib, everolimus, and lenvatinib plus everolimus (see figure 1). A clinical expert explained that there is an evidence gap for treatments at second and subsequent lines. They explained that treatment is decided based on clinical need (using similar criteria as first-line treatment) and which treatments were used before. Clinical experts stated that people with advanced RCC would usually have a maximum of 4 lines of treatment before best supportive care, but most people will only have 1 or 2. Best supportive care consists of monitoring disease progression, symptom control, and end of life care without active treatment. A clinical expert noted that treatment efficacy was expected to diminish with each line and that there was substantial attrition between lines. A clinical expert considered that the treatment pathway modelled over 4 lines of treatment accurately reflected NHS practice. The committee concluded that later-line treatments used in the model represented those used in NHS clinical

practice. See section 2.3.2 of the EAG's assessment report for more details on how first-line treatment was modelled.

Figure 1: Treatment pathway



Abbreviations: ECOG, Eastern Clinical Oncology Group; mTOR, mammalian target of rapamycin; TA, technology appraisal; TKI, tyrosine kinase inhibitor; VEGF, vascular endothelial growth factor; PD1, programmed cell death protein 1; PDL1, programmed death-ligand 1; VEGF, vascular endothelial growth factor.

Notes: PD1/PDL1 use only if no PD1/PDL1 inhibitor in the advanced setting or within 12 months of adjuvant or neoadjuvant treatment. TKI use at second line only if TKIs not had at first line as alone or in combination. TKI use at third line only if TKIs not had at first or second line alone or in combination. Best supportive care consists of monitoring disease progression, symptom control, and end of life care without active treatment.

Economic model structure

1.6 The committee considered the EAG's economic model structure. The EAG constructed a state transition model to estimate the costs and benefits of cabozantinib plus nivolumab and other existing first-line treatment options for RCC (see section 1.4 to 1.5). After first-line treatment, the model captures up to 3 additional lines of subsequent treatment followed by best supportive care. Each line contains an on- and off-treatment health state, and people can enter the best supportive care or death health states at any time. People in the model start on first-line treatment and accrue costs and benefits (quality-adjusted life-years [QALYs]) specific to each health state. See figure 2 for health states and

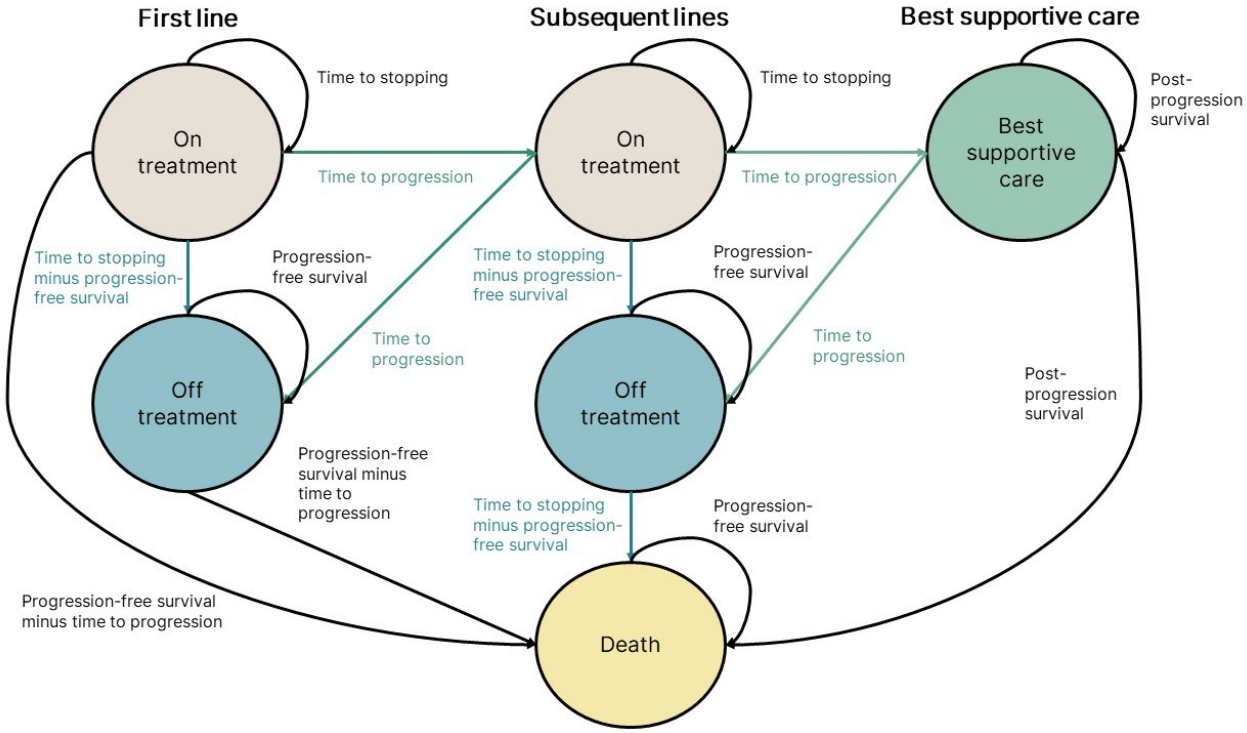
the transitions between them. The model was constructed from an NHS and personal social services perspective over a 40-year time horizon. It has a weekly cycle length and costs and outcomes are discounted at 3.5% per year. The committee also considered the EAG's alternative partitioned survival model approach to investigate structural uncertainty and the interaction between outcomes and any impact on cost effectiveness. The company who manufactures cabozantinib (Ipsen) explained that outcomes generated by the state transition model and the partitioned survival model differed. Ipsen argued that there was value in both approaches, but a partitioned survival model would be more consistent with previous NICE technology appraisals for RCC. The committee noted that the state transition model was more flexible and had a greater ability to explore uncertainties and alternative assumptions across different lines of treatment. The committee explained that the additional flexibility could lead to more uncertainty being presented. The committee noted that this uncertainty would likely be present in the analysis but not visualised when using a simpler partitioned survival modelling approach. The committee would prefer to explore these uncertainties rather than ignore them. The committee preferred the state transition model, because a fundamental part of the pathway model approach is the ability to model multiple lines of treatment in as much detail as possible. But the committee highlighted that it was useful to consider alternative model structures and approaches to investigate relationships between outcomes, especially for instances where surrogacy relationships (relating to correlations between clinical outcomes) break down. For full details on the model structure, see section 4.3.1 of the EAG's assessment report.

Lines of treatment

1.7 Ipsen accepted that the EAG's model was flexible to model up to 4 lines of treatment before best supportive care. But noted that most costs and benefits are accumulated in the first 2 lines of treatment. They considered that the model should only explicitly consider 2 lines of treatment and

argued that this would be consistent with previous NICE technology appraisals in RCC. The EAG explained that only a small proportion of time is spent having third- and fourth-line treatments, which matches expectations for NHS practice. A patient expert commented that the state transition model should consider the entire pathway and that 4 lines of treatment before best supportive care was more appropriate. This will more accurately capture the current treatment pathway, but also ensure the model can change in the future should the pathway change or new treatments become available at third or fourth line. A clinical expert agreed and considered that other technologies could be licensed after second line. So, it would be an advantage to be able to include these in the model efficiently. The committee concluded that, even though most QALYs were accrued in the first and second lines of treatment, it is a strength of the analysis that the model has the option to consider later lines more granularly until the time horizon. The committee preferred the pathway model to include up to 4 lines of treatment followed by best supportive care.

Figure 2: Model structure and transitions between health states



Modelling the treatment effect

Literature review approach

1.8 The committee considered the EAG's sources used to inform clinical effectiveness. Systematic literature reviews were done following NICE's manual on health technology evaluation to identify randomised controlled trials and real-world evidence for advanced RCC. The EAG considered evidence submitted by companies, professional and patient organisations. The committee agreed with the approach and the data sources included to inform the clinical-effectiveness evidence. See section 3 of the EAG's assessment report for more details of the search and selection process and results.

Limitations of included trials

1.9 The committee understood that the clinical evidence base across the risk groups has limitations. For example, some trials only reported the overall population and differences in available treatments, because the pathway has changed over time (newer studies include more novel treatments than older studies). The committee concluded that, despite the limitations, the network meta-analyses included all appropriate data for the pathway model approach, and that the limitations related generally to a changing evidence base. The committee recommended that future RCC trials should be sufficiently powered to analyse differences in treatment effect by risk group (see recommendations for research). See section 3.3 of the assessment report for more details on limitations of the network meta-analyses.

Generalisability

Clear and non-clear cell RCC

1.10 The trials mostly included clear cell RCC, which has better treatment outcomes than non-clear cell RCC. This reduced the applicability of the trial data to other RCC histologies. A clinical expert explained that non-clear cell is normally an exclusion criterion in RCC trials, but can still be

included in the marketing authorisation for treatments. The committee concluded that, without evidence of a differential treatment effect, it was reasonable for the results of the model to be considered generalisable to both clear and non-clear cell RCC, even though trials mostly include clear cell RCC alone. The committee noted that further research on how clear and non-clear cell RCC respond to different treatments would be useful.

Adjuvant pembrolizumab

1.11 The committee noted that all the trials were started before adjuvant treatment with pembrolizumab was available in NHS practice. Clinical experts explained that adjuvant pembrolizumab may improve the prognosis of people with RCC, especially for the favourable-risk group. This is because people have a lower risk of progression and changes in the cancer are more likely to be identified in routine scanning after adjuvant treatment. But there is limited evidence to support this. Because adjuvant pembrolizumab may affect the effectiveness of subsequent treatments, it could make the treatment effect of immunotherapies from trials less relevant to NHS practice. The NHS England Cancer Drugs Fund clinical lead noted that people cannot have immunotherapy treatment in the advanced setting in the NHS if they have had adjuvant pembrolizumab in the last 12 months. They explained that adjuvant pembrolizumab uptake has not been uniform, and they did not expect an increase in adjuvant treatment to affect the proportions of immunotherapies used in the advanced setting. The committee understood that there was insufficient evidence available to model the impact of adjuvant pembrolizumab. It recommended that research could be done to understand the impact adjuvant treatment has on the effectiveness of advanced treatment. The committee concluded that this remains an uncertainty in the model.

Real-world evidence

1.12 The EAG only considered 1 out of 12 real-world datasets identified in the systematic review to be robust and relevant to the UK (Challapalli et al.

2022). The dataset owners gave the EAG access to unpublished patient level data. The data contained IMDC risk scores, treatment patterns, overall survival, progression-free survival, treatment stopping, time to next treatment, time to progression and relative dose intensity. The model uses real-world evidence from Challapalli for the baseline characteristics, natural history, treatment pathway and sequences. Sunitinib data from Challapalli was used to inform a reference curve for the baseline risk for people having first-line treatment. Cabozantinib data was used to inform second-line treatment, and beyond. These reference curves are the foundation of baseline risk in the model. The committee considered that clinical trial populations are often substantially different to people in clinical practice. Most clinical trials recruit people with higher performance status and people have more monitoring than in clinical practice. The committee also noted that nephrectomies are more common in trials than in practice, with about 70% of people in CheckMate 9ER having a nephrectomy compared with 54% in Challapalli. The committee agreed that Challapalli is likely to be representative of people having treatment for RCC in the UK. It explained that Challapalli provides a good source of evidence for baseline characteristics, and agreed it is good practice to use evidence that reflects NHS practice. The committee also explained that the dataset provides a good indication of the likely treatment sequences that the pathway will rely on. Ipsen and other stakeholders raised concerns about transparency and the lack of access to key data because of confidentiality marking. The company highlighted that having no visibility or access to efficacy data or the treatment sequences from the real-world dataset meant that it could not replicate or validate key results. The EAG explained that it made every effort to allow the company to see data, but this was beyond its control. The EAG provided all stakeholders with survival curves fitted to the real-world dataset and results using dummy sequencing data for an executable model. The committee acknowledges that full access would have been ideal, and notes that greater access to the real-world dataset may be possible when the data is

published. See section 3.6 of the EAG's assessment report for more details on the real-world evidence.

Baseline risk

1.13 Sunitinib data from the real-world evidence was used to form the baseline risk for first-line treatment in the model. This data was used to form a 'reference curve'. Cabozantinib data from the real-world evidence was used for the second- and later-line reference curves. A log-logistic model was used to extrapolate first- and second-line progression-free survival reference curves and first-line time to progression curves. A log-normal model was used to extrapolate second-line time to progression curves to cover the time horizon of the model. The committee concluded that the approach and extrapolations used to model baseline risk were appropriate. See section 4.3.5.1 of the EAG's assessment report for more details.

Appropriateness

1.14 Ipsen noted that the real-world evidence did not include outcomes for lenvatinib plus pembrolizumab or cabozantinib plus nivolumab. Both treatments did not feature in the real-world dataset because they are not currently NHS standard care (lenvatinib plus pembrolizumab is a relatively new treatment option because it was recommended by NICE in 2023 when cabozantinib with nivolumab was being appraised). Ipsen noted that there was a lack of detail on how external validity of the real-world evidence was assessed. This made it challenging to assess whether the data was generalisable. A clinical expert explained that the patient population included in the real-world evidence was representative of UK clinical practice. They commented that the proportion of people in each risk group and the number of lines of treatment used were appropriate. They also noted that the proportion who had had a nephrectomy was much more aligned with NHS practice than in the clinical trials (see section 1.12). The committee considered that using the Challapalli dataset was in line with the principles outlined in NICE's real-world evidence

framework and the advice given in the NICE Decision Support Unit Technical Support Document 13. It concluded that despite the limitations with real-world evidence (critiqued in section 3.6.2.4 of the EAG report), the UK real-world dataset reflected NHS practice. The committee further concluded that data used for baseline characteristics, natural history of RCC and treatment sequences were appropriate for the pathway model.

Network meta-analyses

1.15 The trials from the systematic literature reviews (see) were used to inform network meta-analyses for clinical outcomes to be used in the model. Network meta-analyses were done using the first-line networks for the all-risk group, favourable-risk, and intermediate- or poor-risk subgroups. The second and subsequent line network was used for a network meta-analysis only for the all-risk group. Networks were formed for overall survival, progression-free survival, overall response rate, stopping because of adverse events and the risk of treatment-emergent adverse events of grade 3 or higher. The EAG also did a scenario analysis where time to next treatment was used as a proxy for nivolumab plus ipilimumab progression-free survival in the progression-free survival network meta-analysis. This scenario was done to investigate surrogacy assumptions between progression-free survival and overall survival for nivolumab plus ipilimumab. The full rationale is described in section 1.21. The committee concluded the networks were appropriate and considered all relevant outcomes and treatments in the pathway. The committee considered that including nivolumab plus ipilimumab time to next treatment data in the progression-free survival network meta-analysis was imperfect, but provided an additional point of evidence for consideration. See section 3.7 and appendix E of the EAG's assessment report for more details on the network meta-analyses. Following consultation, the EAG updated first-line network meta-analyses with progression-free survival data provided by stakeholders that was previously unavailable for lenvatinib plus pembrolizumab. See section 2.2 of the EAG's response to consultation document for full details of the update to the network meta-analyses.

Relative effectiveness

1.16 The relative effectiveness from network meta-analyses were applied to the reference curves. For first-line treatments, the relative effects calculated in first-line network meta-analyses are applied to the sunitinib reference curves. For second- and later-line treatments, the relative effects calculated in second- and later-line network meta-analyses are applied to the cabozantinib reference curves. The EAG did network meta-analyses for overall survival and progression-free survival for the all-risk group, and the favourable-risk and intermediate- and poor-risk subgroups. The EAG also did a safety network meta-analysis for the all-risk group. The committee acknowledged that network meta-analysis methods were necessary to be able to compare all treatments in the pathway.

First-line relative effects

1.17 The committee considered the EAG's methods for generating relative effects through network meta-analyses. The EAG investigated proportional hazards and fractional polynomial approaches. A proportional hazards approach assumes the relative effects of each treatment compared with sunitinib remains constant over time using hazard ratios. A fractional polynomial approach allows the relative effects to change over time, by generating time-varying hazards. For first-line efficacy, the EAG initially tested the proportional hazards assumption, because proportional hazards network meta-analyses need fewer parameters to be estimated. The proportional hazards assumption was violated for some treatments in the pathway, including nivolumab plus ipilimumab and lenvatinib plus pembrolizumab, as the relative effect compared with sunitinib changed over time. The EAG explained that a violation in one comparison risks carrying through the network and producing implausible hazards for overall survival and progression-free survival for each treatment in the network. So, the EAG used fractional polynomial network meta-analyses to compare progression-free survival and overall survival between first-line treatments in the model base case. Scenarios were provided using the proportional hazards approach. The EAG explained that the fractional

polynomial approach better modelled the observed data by allowing time-varying hazards. Ipsen preferred to use a proportional hazards approach, noting that this simpler approach relied on fewer parameters and was more consistent with previous NICE technology appraisals in RCC. The committee considered this but noted that the proportional hazards assumption was not met for the network. The committee also considered that the flexible time-varying hazard ratios from a fractional polynomial approach provided a better, more plausible fit to observed short-term data. It concluded the fractional polynomial approach was preferred at first line.

Second- and third-line relative effects

1.18 For second and third lines, the model used a proportional hazards approach using the second- and later-line network meta-analysis because of limitations in the data (see section 1.9 and section 1.15 for further details). For fourth-line treatment, a hazard ratio derived from pooled third- and fourth-line outcomes from the UK real-world evidence study was applied to generate a fourth-line curve. Because outcomes were worse at fourth line than third line, this approach effectively 'down-weighted' outcomes at later lines. A clinical expert explained that treatment efficacy is expected to diminish with each line (see section 1.5) and the committee considered that this down-weighting method reflected this. The committee concluded that, while it would have preferred to see a consistent approach applied across all lines, without an available alternative, the proportional hazards network meta-analyses were acceptable to use for subsequent lines.

Limitations

1.19 The committee considered CABOSUN's inclusion in the network meta-analyses. CABOSUN is an older trial, with a small population that included only intermediate- and poor-risk RCC. The committee observed that the progression-free survival and overall survival results for sunitinib reported in CABOSUN were lower than in other trials. It noted that this could have

been because of the population, and because immunotherapies were not available when CABOSUN was done. Ipsen highlighted that the trial only included people with intermediate- or poor-risk cancer, which leads to an overestimation of the treatment effect compared with sunitinib. So, it felt CABOSUN should not be included in the overall network. The committee considered that including CABOSUN in the network meta-analysis could overestimate the relative treatment effect of cabozantinib. The committee concluded that it was cautious about using the CABOSUN trial in the network meta-analyses, which added uncertainty to the results for cabozantinib. To resolve this uncertainty, the EAG assumed equal first-line progression-free survival for cabozantinib and sunitinib for intermediate- or poor-risk disease. The committee was satisfied that this assumption had face validity and provided an additional point of evidence for consideration.

Other modelling assumptions

Surrogacy between outcomes

1.20 Unlike progression-free survival and overall survival outcomes, there was insufficient published trial data on time to progression, time to next treatment and time to stopping treatment to inform standalone networks for these outcomes. The EAG did a targeted review to investigate the plausibility of surrogacy between progression-free survival, time to stopping treatment, and time to next treatment. Based on this review, the EAG applied hazard ratios from the progression-free survival network meta-analysis to the time to stopping treatment and time to next treatment reference curves. It did this to estimate time to stopping treatment and time to next treatment for other treatments. Ipsen considered that there were lots of assumptions involved in generating time to stopping treatment estimates and suggested a simplification in which time to stopping treatment is assumed to be equal to progression-free survival. The committee considered this but noted that, while simpler, assuming that time to stopping treatment was equal to progression-free survival was a strong assumption. The committee considered the evidence and observed

that there was moderate to high correlation between progression-free survival and both time to next treatment and time to stopping treatment for most comparators. It noted that for nivolumab plus ipilimumab the relationship was less clear. The clinical expert explained that time to stopping treatment, time to progression and progression-free survival are not always similar, especially with immunotherapies. This is because some people may stop treatment because of adverse events but may still benefit from the treatment for some time. In these situations, time to stopping treatment will be somewhat shorter than progression-free survival or time to progression. The committee noted that if time to stopping treatment was assumed to be equal to progression-free survival, the off-treatment health states effectively disappeared from the model. It considered that setting time to stopping treatment and time to progression as equal to progression-free survival would bias the results of the model. The committee concluded that it preferred to use available time to stopping treatment data and apply progression-free survival network meta-analyses to the time to stopping treatment and time to progression reference curves. See section 4.3.1.2 of the EAG's assessment report for further details.

1.21 A key assumption of the state transition model is that progression-free survival is an appropriate surrogate for overall survival. This is because the model is driven by multiple lines of progression-free survival to generate survival and quality-adjusted survival outcomes. So the model requires a surrogate relationship between progression-free survival at each line and overall survival to exist. The committee considered that the available evidence in the literature supported the assumption of surrogacy between progression-free survival and overall survival. But the mechanism of action of some treatments meant that the assumption was sometimes limited. For example, nivolumab plus ipilimumab was seen to have worse progression-free survival in CheckMate 214 than other combination treatments in their pivotal trials, but still has a sustained survival benefit. When considering the most recent publicly available data

cut, nivolumab plus ipilimumab had a median progression-free survival of 12.3 months (Motzer et al. 2022) compared with 23.9 months for lenvatinib plus pembrolizumab (see the EAG's assessment report table 14). But, when considering overall survival, this translated to a median overall survival of 55.7 months for nivolumab plus ipilimumab compared with 53.7 months for lenvatinib plus pembrolizumab (see the EAG's assessment report table 13). The EAG explained that this could be caused by tumour flare or pseudoprogression. This is when tumours increase in size in the initial stages of treatment, resulting in a progression event being recorded, before falling in size as the full treatment effect is realised. No evidence of pseudoprogression was identified for nivolumab plus ipilimumab in RCC. But given evidence for this in melanoma and the large difference between observed time to next treatment and progression-free survival in CheckMate 214, the EAG considered it a plausible reason. Alternatively, the potential lack of surrogacy between progression-free survival and overall survival may be because the definition of progression used in CheckMate 214 was different to other trials (investigator assessed compared with independent assessed). Clinical experts explained that pseudoprogression is often discussed when considering immuno-oncology (IO) treatments. They would not expect pseudoprogression to have a major impact on the outcomes for nivolumab plus ipilimumab. They explained that time to next treatment as an outcome is difficult for nivolumab plus ipilimumab because people can get multiple treatment-free intervals when they have not come off treatment entirely and still have benefit before resuming treatment. The clinical experts explained that, because nivolumab plus ipilimumab has a different mechanism of action to the IO-TKI combinations, they would expect outcomes to differ. The experts explained that, because of the differences in modes of action, they expect IO-IO combinations to have worse progression-free survival but better overall survival and IO-TKI combinations to have better progression-free survival, but this would not be translated to similarly sized overall survival gains. The clinical experts acknowledged that it would be difficult to program one model that could

capture the benefits of both combination classes. The committee observed that predictions for overall survival generated by the state transition model for nivolumab plus ipilimumab were more pessimistic than those observed in CheckMate 214 and data from the NHS systemic anticancer therapy (SACT) database. This could have been driven by the breakdown of surrogacy between progression-free survival and overall survival for this technology. This was less of an issue when using the partitioned survival method in scenarios because it uses overall survival data directly. This allows the survival benefit seen in CheckMate 214 to be captured. The committee acknowledged that a partitioned survival modelling approach has limitations compared with a state transition approach. These include reduced flexibility, limited ability to capture later-line costs and benefits, and the need to make other strong assumptions that could lead to additional uncertainty. The EAG also presented a scenario in which time to next treatment was used as a proxy for progression-free survival for nivolumab plus ipilimumab in the progression-free survival network meta-analysis. The EAG argued that, while imperfect, using time to next treatment might better reflect overall survival expected for nivolumab plus ipilimumab given poor surrogacy between progression-free survival and overall survival. The EAG explained that, when time to next treatment is used, the extrapolation fit well to the observed overall survival for nivolumab plus ipilimumab in the real-world evidence. The committee considered that, when there is evidence of poor surrogacy between progression-free survival and overall survival for a treatment in the model, alternative ways of driving health state occupancy should be explored. The committee explained that the EAG time to next treatment scenario was imperfect but provided an additional point of evidence for consideration. It considered that the EAG base case using progression-free survival is likely to underestimate expected overall survival for nivolumab plus ipilimumab. The committee explained that the time to next treatment scenario predicted better overall survival for nivolumab plus ipilimumab, and outcomes more in line with clinical expectations. The committee concluded that overall survival for

nivolumab plus ipilimumab likely fell between the EAG base case and the time to next treatment scenario, and both were important analyses to consider.

Treatment effect waning

1.22 The EAG applied treatment effect waning to the hazards of all IO-TKI combinations at 5 years. This time point was selected based on how long people have these immunotherapies in clinical practice, in which stopping rules are in place. Five years was the longest timepoint when data was available with a reasonable number at risk. The committee agreed that these assumptions were reasonable, but would have preferred to see real world evidence to justify treatment effect waning assumptions. See section 4.3.5.3 of the EAG's assessment report for more details on treatment effect waning.

Sequencing subsequent treatments

1.23 The model includes cost and outcomes for up to 3 lines of subsequent treatment. The model assumes that the type of subsequent treatment is independent of the risk group modelled at first line but is dependent on what treatment was had. Clinical advice and routine commissioning rules were used to determine the plausible sequence after each possible treatment at first, second and third line. Proportions of each treatment observed in the real-world evidence were used to capture subsequent treatments in the model. When a subsequent treatment was implausible, the proportion was set to 0 and the treatment's shares were reweighted across other plausible options. Stakeholders explained they had no access to data and assumptions made for subsequent treatments (see section 1.12). Clinical experts explained that the proportion of people moving on to each treatment at each line in the real-world evidence was plausible and the treatment rules applied were appropriate. Sequences are less certain at later lines, but the committee concluded that the proportions applied to later lines are appropriate. The data did not include lenvatinib plus pembrolizumab or cabozantinib plus nivolumab. Both

treatments did not feature in the real-world dataset because they are not currently NHS standard care. The committee agreed that assumptions used to capture subsequent treatment in the model reflected expected clinical practice. See section 4.3.5.1 of the EAG's assessment report for full details on how clinical effectiveness was modelled for subsequent treatments.

Adverse events

1.24 Checkmate 9ER trial data was used to form the baseline adverse event risk in the model for cabozantinib plus nivolumab and sunitinib at first line. Everolimus data from CheckMate 025 was used for the baseline adverse event risk at second and third line. Hazard ratios from the adverse event proportional hazards grade 3 or more adverse event network meta-analysis were applied to sunitinib data at first line and everolimus data at later lines to estimate adverse events rates for other treatments in the model. Three additional adverse events: hand-foot syndrome, diarrhoea and fatigue were also included, CheckMate 9ER data informed by the baseline rate and relative effects were informed by a Cochrane review. The committee considered that clinical trial data might underestimate the incidence of adverse events compared with clinical practice, but acknowledged that these limitations are a feature of the available data. The committee preferred the network meta-analysis approach to model adverse events over a naive comparison. See section 4.3.6 of the EAG's assessment report for more details on how adverse events were modelled.

Utility values

Health-related quality of life

1.25 The model considered utility values dependent on both progression status and line of treatment. Each line of treatment had a progression free and progressed disease-specific utility, and a utility value for best supportive care. This approach means utility falls as people move through the model and their cancer progresses. The patient experts explained the wide-

ranging effects that a diagnosis of advanced RCC can have on quality of life. They explained that the disease burden and effects of treatment takes a toll on people throughout the course of the condition, resulting in non-trivial decrements in quality of life.

Source of utility values

1.26 Utility values for first-line treatment were sourced from the JAVELIN Renal 101 trial, which was considered appropriate in NICE's technology appraisal guidance on avelumab with axitinib for untreated advanced RCC. The progression-free utility for second line was the same as the progressed disease utility from first. The progressed disease value for second and subsequent lines was calculated using the percentage reduction in utility between progression free and progressed disease observed in the AXIS trial and considered appropriate in NICE's technology appraisal guidance on lenvatinib with everolimus for previously treated advanced RCC. Ipsen preferred to use the utility values observed in the CheckMate 9ER trial, stating that the utility values were consistent with other literature estimates. But the committee thought CheckMate 9ER utility values were implausibly high. It noted that they were broadly consistent with age-matched values for the general population, and this was implausible for people with advanced stage cancer. The committee also heard that the utility values observed in CheckMate 9ER were higher than those accepted in previous RCC appraisals. The committee noted that the small drop in utility from progression-free to progressed disease did not reflect the full expected impact RCC has on quality of life, as described by patient experts. A further scenario was explored, where the percentage drop in utility from progression free to progressed disease from EAG base case utility values is applied to the baseline utility derived from the CheckMate-9ER. This scenario had minimal impact on model outcomes. The committee considered that it would have preferred to have estimates of quality of life from the real-world evidence. But without this, utility values from other published studies with non-trivial decrements at each line of treatment were considered appropriate and supported clinical

and patient expert opinion. See section 4.37 of the EAG's assessment report for full details of the utility approach used in the model.

Costs

Resource use

1.27 Resource use was sourced from published NICE technology appraisals and costs from published sources (NHS reference costs and the Personal Social Services Research Unit), in line with NICE's methods.

- Healthcare resource use costs were applied weekly in the model.
- End of life costs were based on Nuffield trust report applied as a one-off cost upon death.
- Drug and administration frequency were sourced from the summary of product characteristics of each treatment included in the pathway.
- Drug costs were sourced from the BNF or eMIT and confidential discounts were applied, when relevant.
- Proportions of subsequent treatments were informed by real-world evidence and implausible patterns reweighted.
- Subsequent treatment costs were calculated for each line according to the time spent progression free and on treatment in each line.
- Surgery and radiotherapy subsequent treatment costs included in best supportive care are applied as a one-off cost on entering the best supportive care state.
- Adverse event costs were sourced from NHS reference costs and applied as one-off costs to the rates described in section 1.25.

The committee considered resource use assumptions used in the model appropriately reflected NHS practice. The committee was aware that sunitinib costs used in the analysis likely overestimated its true cost, because it has since come off patent. The committee highlighted that these costs should be updated in future analyses. See section 4.3.8 of the EAG's assessment report for more details of the costs and resource use estimates used in the model.

Relative dose intensity

1.28 To accurately capture the cost of treatments to the NHS, the model incorporates relative dose intensities. The relative dose intensities appear lower in clinical practice (reported in the real-world evidence dataset) than those seen in trials. But there were concerns that relative dose intensities collected in the real world are less accurate than those reported in trials. So, the model uses relative dose intensity values collected from the pivotal trial for each treatment included in the model. Clinical feedback explained that there would likely be lower doses used for immunotherapy when used in combination with a high dose TKI, to manage overall toxicity. For lenvatinib, the analysis accounts for the flat pricing structure, in which each tablet of lenvatinib is priced the same regardless of dose. Clinical expert feedback was that most healthcare professionals in the NHS employ a titration phase, in which the dose is gradually increased or decreased over a period of weeks if the person can tolerate the toxicity of their last dose. Lenvatinib is available in 4 mg and 10 mg tablets and has a flat pricing structure. So, as the dose changes the number of tablets needed changes, which has implications on the price. The model accounts for the expected proportion of people that tolerate each dose. The model assumes that:

- 25% of people have 10 mg (1 tablet)
- 57% of people have 14 mg (2 tablets)
- and 18% of people have 20 mg (2 tablets).

Clinical experts agreed that the proportions used in the model were reasonable. The committee was satisfied that the approach taken to reflect changes in dosing regimens aligned with expectations for clinical practice. A further scenario was done, in which, on average, people have a dose of lenvatinib satisfied by 2 tablets. The committee heard that this was likely a pragmatic scenario because:

- the recommended dose for lenvatinib in combination with pembrolizumab for untreated renal cell carcinoma is 20 mg (2 tablets)

- any off-label deviations from the recommended dose (for example, to 18 mg) in clinical practice would likely mean the average number of tablets would likely be 2 per person.

The committee considered both the base-case proportions and the pragmatic 2-pill scenario in its decision making.

Other considerations

Severity

1.29 The committee considered the severity of the condition (the future health lost by people living with the condition and having standard care in the NHS). The committee may apply a greater weight to QALYs (a severity modifier) if technologies are indicated for conditions with a high degree of severity (a severity modifier). The committee considered absolute and proportional QALY shortfall estimates in line with NICE's manual on health technology evaluation. It noted that the severity of the condition depends on which treatment was considered standard care and there are a range of treatments recommended for untreated advanced RCC. The committee was presented with 3 options for assessing whether a severity weighting applied. These were fully incremental analyses, pairwise analyses (in which the most appropriate comparators were defined) and a weighted market share approach. The committee considered that a fully incremental analysis would be the most suited to optimising the treatment pathway, but it recognised that technology appraisals recommend new treatment as 'options'. So, for some technology appraisals, a pairwise analysis could be appropriate if there are defined clinical reasons why specific comparisons should be made, or to consider comparators that are likely to be displaced. The committee explained that all judgements on severity were based on the current model base case and that other analyses made as the pathway evolves and the model is developed may result in different severity conclusions.

Summary of committee's preferred assumptions

Table 1: Summary of committee preferred assumptions

Category	Type	Committee preferred assumptions	Report section
Setting	Perspective	NHS and personal social services	1.6
Setting	Time horizon and cycle length	40-year time horizon and weekly cycle length	1.6
Setting	Discounting	Costs and outcomes discounted at 3.5% per year	1.6
Setting	Model structure	Hybrid state transition model considering 5 lines, split by on- and off-treatment status (up to 4 active treatments followed by best supportive care)	1.7 and 1.8
Setting	Health state transitions	Transitions between lines driven by progression status Transitions between the on- and off-treatment states driven by time to stopping treatment	1.7
Input	Baseline characteristics	UK real-world evidence population	1.13
Input	Effectiveness data	UK real-world evidence	1.13
Input	Adverse events	Grade 3+ adverse event rates in less than 5% of people in CheckMate 9ER for cabozantinib with nivolumab and sunitinib Additional adverse events of interest (hand-foot syndrome, diarrhoea and fatigue) included on clinical advice Safety proportional hazards network meta-analysis applied to reference sunitinib data for non-reference treatments	1.26
Input	Utilities	Utility differs by progression status and line of therapy Use published utility values accepted in previous NICE technology appraisals TA645 to inform first-line progression free and progressed disease utility and second-line progression fee utility, TA498 percentage reduction applied for later lines	1.28
Input	Resource use	Based on previous NICE technology appraisals, supported by clinical opinion	1.29
Input	Subsequent treatments	Proportions sourced from UK real-world evidence	1.25

Input	Costs	Sourced from published sources (NHS Reference costs, PSSRU, Nuffield Trust, BNF, eMIT)	1.29
Input	Relative dose intensity	Calculated using CheckMate 9ER data for cabozantinib plus nivolumab and sunitinib, and published sources for other treatments	1.30
Assumption (efficacy)	Reference treatments	Sunitinib first-line reference treatment as central node in first-line network Everolimus second- and later-line reference treatment as central node in second- and later-line network	1.14
Assumption (efficacy)	Reference extrapolations	UK real-world evidence used to model relevant outcomes at each line for the reference treatment Log-logistic model used to extrapolate progression-free survival and time to progression	1.14
Assumption (efficacy)	Comparative effectiveness	Fractional polynomial network meta-analysis to generate first-line outcomes for non-reference treatments Cabozantinib intermediate- or poor-risk cancer progression-free survival assumed equivalent to sunitinib Proportional hazards network meta-analysis applied to second- and later-line reference outcomes	1.17 to 1.21
Assumption (efficacy)	Surrogacy	Time to stopping treatment and time to progression assumed appropriate surrogates to progression-free survival Hazard ratios from progression-free survival network meta-analysis applied to time to stopping treatment and time to progression outcomes Model driven by progression-free survival, assumes that progression-free survival at each line is an appropriate surrogate for overall survival In the case that surrogacy relationships break down, considering other outcomes is appropriate. The use of time to next treatment as a surrogate for nivolumab plus ipilimumab progression-free survival at first line was appropriate to consider.	1.22 and 1.23
Assumption (efficacy)	Fourth line and PPS	Hazard ratio calculated from third-line and fourth-line real-world evidence outcomes used to calculate survival in fourth line Log-normal model used to extrapolate	1.19

Assumption (efficacy)	Treatment effect waning	Applied at 5 years to immunotherapy and TKI combinations based on hazards, for all endpoints	1.24
Assumption (other)	Treatment sequencing	Treatment rules limit available later lines treatments based on what people have at earlier lines UK real-world evidence proportions of implausible treatments reweighted to other plausible options	1.25
Assumption (other)	Severity	Considered absolute and proportionate QALY shortfall estimates using fully incremental, weighted and pairwise approaches	1.32

2 Recommendations for research

2.1 The committee recommended that future renal cell carcinoma trials should be sufficiently powered to analyse differences in treatment effect by risk group (see section 1.9).

2.2 The committee highlighted areas in the analysis and evidence that could benefit from future research:

- understanding long-term health-related quality of life in the real world (see section 1.26)
- survival benefit of immunotherapies in the real world (see section 1.23)
- understanding how clear and non-clear cell RCC responds to different treatments (see section 1.10)
- the impact of adjuvant treatment on the effectiveness of advanced treatment (see section 1.11).

3 Evaluation committee members and NICE project team

Evaluation committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This model report was considered by committee B. Committee members from committee A, committee C and committee D also took part in the meeting.

Chair

Dr Charles Crawley

Chair, technology appraisal committee B

ISBN: [to be added at publication]