NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

Belzutifan for treating tumours associated with von Hippel-Lindau disease

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using belzutifan in the NHS in England. The evaluation committee has considered the evidence submitted by the company and the views of non-company stakeholders, clinical experts and patient experts.

This document has been prepared for consultation with the stakeholders. It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from the stakeholders for this evaluation and the public. This document should be read along with the evidence (see the committee papers).

The evaluation committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of age, disability, gender reassignment, pregnancy and maternity, race, religion or belief, sex or sexual orientation?

Issue date: [November 2023]

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

After consultation:

- The evaluation committee will meet again to consider the evidence, this evaluation consultation document and comments from the stakeholders.
- At that meeting, the committee will also consider comments made by people who are not stakeholders.
- After considering these comments, the committee will prepare the final draft guidance.
- Subject to any appeal by stakeholders, the final draft guidance may be used as the basis for NICE's guidance on using belzutifan in the NHS in England.

For further details, see NICE's manual on health technology evaluation.

The key dates for this evaluation are:

- Closing date for comments: 03 January 2024
- Second evaluation committee meeting: 15 February 2023
- Details of membership of the evaluation committee are given in section 4

1 Recommendations

1.1 Belzutifan is not recommended, within its marketing authorisation, for treating von Hippel-Lindau (VHL) disease in adults:

- who need treatment for VHL-associated renal cell carcinoma, central nervous system hemangioblastomas or pancreatic neuroendocrine tumours, and
- when localised procedures are unsuitable or undesirable.

1.2 This recommendation is not intended to affect treatment with belzutifan that was started in the NHS before this guidance was published. People having treatment outside this recommendation may continue without change to the funding arrangements in place for them before this guidance was published, until they and their NHS clinician consider it appropriate to stop.

Why the committee made these recommendations

VHL disease is a genetic condition that severely affects the quality of life of people with it, and their families and carers. The condition increases risk of certain tumours developing, and surgery is the main treatment option. There are no licensed medicines for the underlying causes of VHL.

Clinical-effectiveness evidence from a small study suggests that belzutifan reduces tumour size. It also suggests that it increases the amount of time people have before their condition gets worse, but by how much is uncertain.

There are also uncertainties in the economic model, as well as assumptions that likely favour belzutifan. So, it is not clear what the most likely cost-effectiveness estimates are for belzutifan and it cannot be recommended for routine use.

Even though some of the uncertainty in the clinical-effectiveness evidence could be addressed in the Cancer Drugs Fund, the cost-effectiveness evidence suggests that

belzutifan is not likely to be cost-effective. So, belzutifan is not recommended for use in the Cancer Drugs Fund.

2 Information about belzutifan

Marketing authorisation indication

2.1 Belzutifan (Welireg, Merck Sharp & Dohme) is indicated for 'the treatment of adult patients with von Hippel-Lindau (VHL) disease who require therapy for VHL-associated renal cell carcinoma (RCC), central nervous system (CNS) hemangioblastomas, or pancreatic neuroendocrine tumours (pNET), and for whom localised procedures are unsuitable or undesirable'.

Dosage in the marketing authorisation

2.2 The dosage schedule is available in the summary of product characteristics for belzutifan.

Price

2.3 The list price for a 90-tablet pack of 40 mg belzutifan is £11,936.70 (excluding VAT; BNF online, accessed November 2023).

2.4 The company has a commercial arrangement, which would have applied if belzutifan had been recommended.

3 Committee discussion

The evaluation committee considered evidence submitted by Merck Sharp & Dohme, a review of this submission by the external assessment group (EAG), and responses from stakeholders. See the committee papers for full details of the evidence.

Clinical management

The condition

3.1 Von Hippel-Lindau disease (from now, VHL) is caused by a mutation in the VHL gene. This gene is responsible for producing a protein that controls cell growth. A mutation in the gene can cause cells to grow abnormally, leading to cysts or tumours developing in different parts of the body, such as the kidneys, brain and pancreas. This can lead to renal cell carcinoma (RCC), central nervous system haemangioblastomas (CNS Hbs) and pancreatic neuroendocrine tumours (pNETs). The patient experts explained that the experience of living with VHL varies from person to person. Some people with the condition might only develop 1 or a few tumours in their whole life, while others might have multiple tumours in different organs. There are also a wide variety of debilitating symptoms depending on tumour sites. These include constant pain, loss of balance and motor skills, loss of vision, breathlessness, coughing, headaches, confusion, severe nausea and fatigue. Scans before appointments can cause anxiety and people also worry about disability caused by surgery. The clinical experts stated that, with more effective treatment, there is potential for people with the condition to live longer and have a better quality of life. Also, the patient experts' statement highlighted that caring for a family with VHL is emotionally challenging and has a psychological effect. The patient experts explained that carers are often the family members of people with VHL. They worry that they are not able to help, and have to live with the constant worry that they also carry the gene for VHL and may pass it on to their children and grandchildren. The committee understood that people with VHL often have difficulty doing day-to-day tasks, and fear surgery. Also, the condition can have a negative effect on self-esteem and cause relationship difficulties. Survival with VHL has improved over time, and has become closer to that of siblings without VHL and the general population. Life expectancy for men with VHL is about 67 years and for women is about 60 years. The committee noted that living with the condition and caring for people with VHL is physically and emotionally challenging. It concluded that VHL is a

highly heterogeneous condition, and has considerable physical and emotional effects from repeated surgeries and anxiety from scans.

Unmet need

3.2 Surgery and other localised procedures are the main treatment options for people with VHL, but sometimes they are not appropriate. The clinical and patient experts explained that there is an unmet need for effective new treatments for people with VHL when surgery is unsuitable. They explained that people often have repeated surgeries throughout their life, and this is the only way to remove VHL tumours. Because of this, people may lose their eyes, a part of or a whole kidney, or their pancreas. They may also develop neurological issues such as paralysis after spine or brain surgery or may need lifelong medical intervention such as dialysis. The patient experts explained that avoiding multiple surgeries could greatly improve physical and mental wellbeing, and improve quality of life for both people with VHL and their carers. The committee noted that there is an unmet need for treatments that could improve outcomes and quality of life for people with VHL. It concluded that the current unmet need could be addressed by belzutifan because it has the potential to preserve or delay the loss of organ function and the associated morbidity.

Existing treatment

3.3 The patient and clinical experts explained that there is no treatment that addresses the underlying cause of VHL. It is a heterogeneous condition and clinical management usually is done by a multidisciplinary team. They explained that VHL management starts with diagnosis through genetic counselling or de novo testing. This is followed by regular surveillance by genetics services. Surveillance aims to detect tumours and monitor tumour growth. MRI or ultrasound examinations of the abdomen are done every 12 months for RCC tumours and pNETs. For CNS hemangioblastomas, MRI scans of the head are done every 12 to 36 months. If VHL-associated tumours need treatment, interventions are

generally based on a specific threshold. The clinical experts explained that RCC tumours are kept under surveillance until they reach a diameter of 3 cm, and pNETs until they have a diameter of 2 cm. At this stage, the risk of metastasis exceeds the benefit of organ preservation and surgery can be recommended. Surgery for CNS Hb is normally needed when the tumours grow to a size that causes symptoms. The patient experts explained that standard care varies, and focuses on preventing tumour growth and metastasis while preserving organ function. The clinical experts explained that surgery is highly effective in most cases for all VHL tumours, with the most benefit for VHL-associated RCC. But surgery can result in organ loss after multiple surgeries or morbidity, depending upon the primary VHL-associated tumour. The clinical experts also explained that some people have VHL tumours that have grown to an extent that localised procedures are unsuitable or undesirable. This might be because of an increased risk of loss of organ function or risks from the procedures. These include people with:

- RCC: expected surgery that would result in bilateral nephrectomy and the need for dialysis
- pNETs: expected surgery that would result in pancreatectomy (or pancreaticoduodenectomy, also known as the Whipple surgery) and that may lead to long-term morbidities such as diabetes
- CNS Hb: expected surgery with a high risk of complications (such as tumours at the craniospinal junction) that could result in neurological deficit or paralysis.

The committee noted that managing VHL is highly complex, heterogeneous and individualised. It noted that the decision to have surgery, or to delay surgery and have active surveillance, is based on several factors. These factors include the type of tumour, its location, the individual's overall health, suitability and need for surgery. The committee also understood that belzutifan should only be prescribed if the clinician

and person at risk agree that the benefits of treatment and the potential delay of surgery outweigh the risks.

Belzutifan marketing authorisation and positioning

3.4 The committee noted that the indication for belzutifan (see section 2.1) involves some subjectivity, especially for 2 criteria:

- when treatment is needed
- when localised procedures are unsuitable or undesirable.

The committee noted that this introduces challenges in clearly defining the population who should be eligible for belzutifan treatment, and at which stage of VHL. The company clarified the definition of its indication and positioning in the treatment pathway. It explained that:

- when treatment is needed refers to a need for surgery or a related procedure when tumours reach a certain size (for RCC tumours, a more than 3 cm diameter; for pNETs, a more than 2 cm diameter; for CNS tumours, that they cause symptoms)
- unsuitable or undesirable refers to when localised procedures (surgery and radiotherapy procedures) would result in organ loss or severe functional deficits (see section 3.3).

The clinical experts clarified that people who are fit enough and whose VHL is eligible for surgery would normally have the surgery because it is an effective option. They also explained that belzutifan will provide an option for people when a tumour reaches the treatment threshold and they are waiting for the surgery. The patient and clinical experts explained that this would be at about 4 months. The experts also explained that, in clinical practice, surgeries are not done immediately except for rare cases of certain CNS Hbs. This is because people and their tumours are monitored closely over time. The committee noted that there may be intervals of 4 months between a tumour reaching the treatment threshold and the decision to proceed with surgery. It also noted that belzutifan is

positioned for when people have had surgery for tumours that reached the treatment threshold and are having active surveillance until the next surgery that could lead to potential organ loss is needed. The committee noted that the company's marketing authorisation and positioning of belzutifan in the treatment pathway is subject to interpretation. This may be challenging to implement in clinical practice.

Relevant population

3.5 The company's trial population was at a different position in the treatment pathway than that in NICE's final scope on belzutifan for treating tumours associated with VHL disease and the marketing authorisation. The marketing authorisation population consists of adults with VHL who need treatment for VHL-associated RCC, CNS Hb, or pNETs, and for whom localised procedures are unsuitable or undesirable. In contrast, the company's main clinical trial (MK-6482-004) included people who:

- had at least 1 more measurable VHL-associated RCC only (could have other tumours), and
- did not need imminent surgery and may have had pNETs or CNS Hbs.

The company compared belzutifan with standard care. This comprised surgery and active surveillance for RCC, CNS Hb and pNET cohorts. The EAG noted that MK-6482-004 excluded people who had an immediate need for surgery for tumours and for whom localised procedures were unsuitable or undesirable. It explained that this meant there was a misalignment between the marketing authorisation and MK-6482-004. The EAG further explained that there could be clinical differences between the trial, marketing authorisation and comparator populations. That is, people for whom surgery is deemed suitable (the comparator) are likely to be fitter than the marketing authorisation population. It also explained that people for whom surgery is deemed to be needed may have a greater tumour burden than people recruited to MK-6482-004. The committee noted that MK-6482-004 represented a population with different needs

than the population in the marketing authorisation. It also noted that MK-6482-004 was originally designed only to include people with VHL-associated RCC tumours. The occurrence of CNS Hbs and pNETs within the trial population was coincidental, but was granted marketing authorisation for RCC, CNS Hbs and pNETs. The committee recognised that designing a trial in the population of interest would have been difficult for ethical reasons. But it thought that this issue severely limited the generalisability and applicability of the clinical-effectiveness evidence. So, the committee was cautious in interpreting the results from MK-6482-004.

Clinical-effectiveness evidence

The MK-6482-004 trial

3.6 The clinical-effectiveness evidence for belzutifan came from MK-6482-004. This was a multicentre single-arm open-label phase 2 study. It included 61 people with VHL with at least 1 measurable RCC tumour. Fifty of them also had CNS Hbs and 22 also had pNETs. The primary outcome of MK-6482-004 was the objective response rate (complete response or partial response). The secondary outcomes were disease control rate, duration of response, time to response, progression-free survival and time to surgery. At the latest data cut (April 2022), the objective response rate was 63.9% for RCC (95% confidence interval [CI]: 50.6% to 75.8%), 44% for CNS Hbs (95% CI: 30.0% to 58.7%) and 90.9% for pNETs (95% CI:70.8% to 98.9%). This was assessed using the Response Evaluation Criteria in Solid Tumours 1.1 criteria. Disease control rate (complete response, partial response or stable disease) was 98.4% for RCC (95% CI: 91.2% to 100.0%), 90.0% for CNS Hbs (95% CI: 78.2% to 96.7%) and 100% for pNETs (95% CI: 84.6% to 100.0%). The median time to response for RCC was 11.1 months. Progression-free survival results are considered confidential by the company and cannot be reported here. The committee concluded that belzutifan was likely to be clinically effective in reducing tumour size and so the need for surgery. It

noted that there was some uncertainty about how tumour size relates to symptom burden in CNS Hb. The committee also noted the time needed for the response was 11.1 months for RCC. It considered this in terms of the positioning of belzutifan and that assessing whether surgery is needed may be challenging to predict in clinical practice.

Comparator data (VHL natural history study)

3.7 Because MK-6482-004 was a single-arm study, the company used data from a VHL natural history study to inform the comparative effectiveness of the standard care comparator. This study was a retrospective non-interventional study of existing medical records with supplemental electronic medical record data abstraction and review of abdominal imaging scans done during routine clinical care in a cohort of people in the US. It included people with at least 1 VHL-associated RCC tumour measured during the study period. They also had to meet other VHL natural history study eligibility criteria that were identified and followed until the end of the assessment window (July 31 2004 to June 30 2020). The EAG noted that comparative effectiveness results derived from VHL natural study data did not:

- represent the population in the belzutifan's marketing authorisation
- collect the appropriate data to address the population of interest.

The committee noted that the lack of comparator data meant that the company used a matching adjusted indirect comparison (MAIC) method to compare belzutifan with the standard care (see section 3.9). The committee noted that the VHL natural history study was well conducted but was US based. So, it included a population that was narrower than the marketing authorisation population for belzutifan (see section 3.5). The committee concluded that the VHL natural history study was not aligned with the decision problem.

Outcomes

3.8 The company focused on collecting outcomes such as time to response, progression-free survival, time to surgery and overall survival. MK-6482-004 and the VHL natural history study were used to compare the outcomes of treatment with belzutifan with the outcomes in standard care to inform the model. The committee noted that these outcomes are not the same as those used in standard NICE cancer topic evaluations. The committee noted that time to surgery (a key model transition) represented a highly heterogeneous outcome that depended on several factors, such as:

- size and location of tumours
- symptom development
- the overall health of the person with VHL.

It also noted that it is closely related to loss of organ or neurological function, which were not well characterised in the VHL natural history study. The committee concluded that there was considerable uncertainty in eligibility criteria for MK-6482-004 and the VHL natural history study compared with the population of interest and more information may be required on outcomes that more closely match loss of organ or neurological function.

Establishing relative treatment effect

3.9 The relative treatment effects of belzutifan compared with standard care were derived from an indirect treatment comparison (ITC) using the propensity-score weighting-based MAIC methods. This used individual patient data from the VHL natural history study to match baseline characteristics with the MK-6482-004 trial. After matching, this made it easy to compare outcomes because the populations were likely to be more balanced. The company compared the results of MK-6482-004 and collected data for the standard care arm of the cost-effectiveness model. It did this by selecting a subgroup population from the VHL natural history study that matched MK-6482-004's inclusion and exclusion criteria. The

committee noted that, for the comparison with standard care, the company did a series of adjustments, specifically:

- Kaplan–Meier curves were fitted to the VHL natural history study (standard care) and the MK-6482-004 trial data.
- The fitted Kaplan–Meier curves from the VHL natural history study data were then adjusted using MAIC to match the population in MK-6482-004 based on variables such as age, gender, previous surgeries and tumour size.
- Time to surgery, second surgery and metastasis in the VHL natural history study were adjusted to reflect a less active surveillance. This was because the company considered that the standard care seen in the VHL natural history study cohort may have been better than that routinely provided in UK clinical practice.
- An additional assumption that 90% of people with RCC or pNETs, and 50% of people with CNS Hbs, have immediate surgery was then applied.

The committee noted that the relative treatment effect was highly uncertain because of the assumptions needed to convert from MK-6482-004's population to the marketing authorisation population. It considered that the assumption that 90% of people having standard care would proceed to immediate surgery made the MAIC adjustment of the Kaplan–Meier curve relatively unimportant. This was because it only applied the residual 10% of the standard care population. The committee also considered this immediate surgery assumption too simplistic and not evidence based. It noted that the adjustment for the population was only applied to the standard care arm, meaning that the generalisability of time to surgery in a less severe population was implicitly assumed. The committee also noted that the company assumed that almost everyone in the standard care arm would have immediate surgery leading to functional organ loss. But it used the evidence from MK 6482-004, which excluded

such people. The committee understood that this would have substantially bias the comparison in belzutifan's favour. It concluded that it would also like to have seen alternative methods explored such as a simulated treatment comparison in line with NICE DSU Technical Support document 18.

Economic model

Company's model structure and outputs

3.10 The company used a Markov model structure to estimate the cost effectiveness of belzutifan compared with standard care. The model included 5 health states: presurgery, surgery, event-free after surgery, metastatic disease (pre progression and post progression) and death. The model had a lifetime horizon (59 years) and weekly cycle length. In the company's model, people started at age 41 years in the presurgery health state reflecting the treatment decision point. From there, they could transition to surgery, metastatic disease or death health states. The committee questioned the company for using a starting age of 41 years. This was because the marketing authorisation population was likely to have had many more surgeries and was much later in the treatment pathway. The EAG considered the company's model structure was appropriate for only the RCC cohort, but it noted a high level of uncertainty, specifically:

- There was an overlap in the data informing input parameters for the 3 cohorts.
- The rate of surgeries (moving from presurgery to surgery) was based on surgeries for the primary tumour. But it was not clear from the trial and the VHL natural history study whether the data used to specify these rates related to the treatment of primary tumour.
- Including people for whom surgery is not suitable but who would have it immediately as a 'last resort' in the standard care arm only was not appropriate.

The committee noted that the company assumed 90% of people with RCC in the standard care arm had immediate surgery. It also assumed that, in 80% of these people, the surgery would lead to end-stage renal disease or dialysis. The committee considered that the company's model was too complex to include all VHL cohorts. It also thought that the model input parameters were not generalisable to the marketing authorisation population (see section 3.5). The committee considered that the model included 3 cohorts with overlapping evidence and that the assumptions used in the model were not based on firm evidence. It would have preferred to see a model structure based on the natural history of VHL disease rather than individual tumours and the surgery associated with them. It noted that the relative efficacy derived from the ITC (see section 3.10) introduced additional uncertainty along with the uncertainties in modelling assumptions. The committee concluded that the model outputs as a consequence were unreliable for decision making.

Time on treatment

3.11 The company assumed that people would stay on belzutifan until VHL progressed or until they had side effects. It used time on treatment data from MK-6482-004 to model time on treatment with belzutifan. The committee noted that almost half of the people who stopped belzutifan in the trial were reported to have done so through choice. A minority who were reported as stopping stated reasons of progression or side effects. The company explored different parametric fits to the patient level time on treatment data (exponential, Gompertz, log-logistic, log-normal, gamma generalised gamma and Weibull) using patient-level data from MK-6482-004. The company preferred using the Gompertz-based model that was based on statistical fit, visual inspection and clinical relevance in its base case. The company also explored the effect of using the second-best (Weibull) model in the scenario analysis. The EAG explained that there was uncertainty in the long-term extrapolations of this data. The committee noted that it was unclear why a high proportion of people

stopped treatment for the stated reason of choice. It thought that this may not be generalisable to the population of interest in the company's model. This was because the alternative for people stopping treatment would be surgery resulting in organ loss, as assumed by the company in its model. So, the committee concluded it would have preferred to see the modelling using belzutifan continued until progression or until side effects because this would more closely match the target population.

Treatment waning

3.12 The committed noted that, because of uncertainties in the modelled time on treatment, the company's model incorporated a treatment effect waning using 'off-treatment' health states. Transition probabilities from the 'off-treatment' health states were assumed to gradually converge to the respective values in the standard care arm based on data from the VHL natural history study. The company assumed treatment waning occurs gradually over 2.71 years (the amount of time until the largest RCC tumour reaches the baseline levels of growth) from the end of the maximum follow-up (3.84 years). It also assumed that tumour growth would return to baseline level at an average rate of 3.52 mm per year after stopping. The same assumption was used for CNS Hbs and pNETs. The EAG explained that the duration of assumed residual benefit over a period of 2.71 years might be appropriate for RCC but could be different for CNS Hbs and pNETs. The committee noted that treatment waning might be appropriate because of the tumour size reduction seen in MK-6482-004. But it reiterated that there were concerns around the generalisability of the data for the population of interest. The committee concluded that it would like to have seen extensive sensitivity analyses and testing alternative assumptions on the treatment waning of effect across the tumour types.

Health-related quality of life

3.13 In MK-6482-004, no health-related quality-of-life data was collected. The committee noted that the company used health-state utility values in the

economic model. These were derived from the VHL real-world quality-of-life disease burden study and a trial of pembrolizumab in the adjuvant treatment of RCC (KEYNOTE-568). It considered the utility values to be mostly appropriate but noted that these were not measured in the population of interest. The company's model also applied disutilities from tumour-associated surgeries and surgical complications (including loss of organ function). The company used a variety of sources to derive disutility values for long-term and short-term complications, including the VHL real-world quality-of-life disease burden study and published literature. The committee noted that the company had used an additive approach to adjust disutility values in the model. It considered that the disutility values adopted by the company for long-term complications of surgery may have lacked face validity. The committee also noted that the multiplicative method is a preferred approach for combining disutilities, in line with NICE's health technology evaluations manual. It considered that the outputs of the effective utility loss for people after surgery were uncertain. The committee was aware that some of the utility decrements the company used to represent the long-term consequences of surgery were large. It noted that people having dialysis had a utility value of -0.527. This is much larger than the estimates seen in the literature and the values used in NICE's guidelines on renal replacement therapy and conservative management and chronic kidney disease. The committee understood that the company had derived the disutility values by deducting estimates of quality of life on dialysis from 1. The committee agreed that this approach was not appropriate because it effectively assumes that, if people were not on dialysis, they would be in perfect health. It noted that some of the health-related quality-of-life impact was double-counted, for example, separate disutilities for 'stroke' and 'neurological complications' were applied in the model. The committee considered that there was a more appropriate approach would be to calculate a relative utility impact. This would be done by comparing an absolute estimate of utility on dialysis with an age- and sex-matched expectation from the general population.

The committee concluded that the total effective utility loss after surgery should have been calculated using a multiplicative method. It added that this should have been validated against literature for similar outcomes (such as end-stage renal disease, post pancreatectomy and neurological damage). It should also have taken account of any additional consideration for people with VHL.

Severity

3.14 The committee may apply a greater weight to quality-adjusted life years (QALYs; a severity modifier) if technologies are indicated for conditions with a high degree of severity. So, the committee considered the severity of VHL, that is, the future health lost by people living with the condition and having standard care in the NHS. The company provided absolute and proportional QALY shortfall estimates for all VHL cohorts in line with NICE's health technology evaluations manual. In its analyses, the proportional QALY shortfall was above 0.95 for RCC and CNS Hbs and below 0.95 for pNETs. So, a severity weight of 1.7 was applicable for RCC and CNS Hb and of 1.2 for pNETs. The company explained that in clinical practice or the real world, not all VHL cohorts are distinct from each other. People in MK-6482-004 had more than 1 tumour manifestation (see section 3.5), and pNET is associated with high mortality and morbidity because of pancreatic surgery. So, the company applied a QALY weighting of 1.7 to all 3 VHL cohorts. Based on the QALYs generated from the company's model, the company and EAG agreed that, for RCC and CNS Hbs, the QALY weighting 1.7 was applicable. But the EAG considered it inappropriate to apply the QALY weighting of 1.7 for pNETs. The committee acknowledged the substantial impact of VHL. The committee also noted that both the company's and the EAG's analyses were subject to a high degree of uncertainty because of the underlying assumptions adopted for modelling standard care. The committee concluded it was unable to apply an appropriate severity

weight based on the calculations presented by the company because of uncertainty in its underlying assumptions.

Cost-effectiveness estimates

Uncertainties in the evidence and the company's modelling assumptions

3.15 The committee noted that the company had not provided data on the effectiveness of belzutifan for the marketing authorisation population (see section 3.5). The committee recalled that there were issues with the ITC and the company's model, which needed resolving before it could be considered suitable for decision making. The committee was aware that the EAG was unable to define its base case because of uncertainties in the evidence and assumptions made in the model. In particular, the committee noted the high level of uncertainty in the following areas, in which it would like to see further analyses and exploration by the company:

- the evidence for the generalisability of the MK-6482-004 population to the marketing authorisation population (see section 3.5)
- the ITC approach and using the propensity-score weighting method, which were highly uncertain, with alternative methods explored (see section 3.9)
- the uncertainty in the model input parameters and assumptions and uncertainties in the model outputs (see section 3.10)
- modelled time on treatment for belzutifan until progression or until side effects (see section 3.11)
- extensive sensitivity analyses and testing alternative assumptions on the treatment waning effect across the tumour types (see section 3.12)
- the uncertainty in the company's approach to surgery-associated disutility values, with an exploration of the multiplicative approach and use of validated disutility values against literature for similar outcomes (see section 3.13).

The committee noted that the company's incremental cost-effectiveness ratios (ICERs) for the comparison with standard care were within the range normally considered an effective use of NHS resources. But, because the clinical and economic evidence was too uncertain, the committee concluded that there was no plausible cost-effectiveness estimate on which to base a decision. So, belzutifan could not be recommended for routine use in the NHS.

Cancer Drugs fund

3.16 Having concluded that belzutifan could not be recommended for routine use in the NHS, the committee then considered whether it could be recommended within the Cancer Drugs Fund. The committee discussed the arrangements for the Cancer Drugs Fund agreed by NICE and NHS England in 2016, noting NICE's Cancer Drugs Fund technology appraisal process and methods guide (addendum). The committee recognised that people with VHL have a high unmet need, and the availability of new treatment for VHL is very important. The committee was aware that the company had expressed an interest in being considered for funding through the Cancer Drugs Fund in its submission. This was because of acknowledged uncertainties and lack of data directly relevant to the decision problem. The committee recognised that belzutifan is innovative and its use in clinical practice would help address unmet needs. But it also noted that it had not been presented with sufficient evidence of the plausible potential for belzutifan to be considered a cost-effective use of NHS resources. This was because of considerable uncertainties relating to relative treatment effectiveness and assumptions used in the model. The committee also noted that the Cancer Drugs Fund may provide the opportunity to collect additional data to address some uncertainties about belzutifan's efficacy in clinical practice. But it thought that whether this would provide data to address the issues of its comparative effectiveness compared with standard care was uncertain. So, the committee concluded

that belzutifan did not meet the criteria to be considered for inclusion in the Cancer Drugs Fund.

Other factors

3.17 Because of the rarity of VHL, the committee recognised difficulties in the ability to collect or generate clinical evidence of belzutifan's comparative effectiveness and the natural history of VHL contributing to significant uncertainty in the decision making. The committee also noted that there may be other factors not included in the analysis, such as the potential of belzutifan to reduce fear and anxiety from frequent scans. It considered it could potentially apply greater flexibility in decisions around acceptance of uncertainty and consideration of benefits outside of the ICER calculation in these circumstances. However, it considered additional analysis were possible to reduce these uncertainties (see section 3.15) and that the decision uncertainty outweighed these considerations.

Equalities

3.18 The committee noted that, because VHL is a genetic condition, some families are disproportionately affected. The condition can affect people when they are very young. It also noted that people from deprived areas, with language, learning or cultural barriers, or those with disabilities may be at a disadvantage. The committee agreed that, if belzutifan were recommended, the recommendation would not restrict access for some people over others. No other equality or social value judgment issues were identified.

Innovation

3.19 The company considered belzutifan to be innovative for treating VHL in people with a very high unmet need. The company explained that belzutifan got regulatory approval through the Medicines and Healthcare Products Regulatory Agency Innovative Licensing and Access Pathway, which is reserved for innovative medicines. The clinical experts also

considered belzutifan a step change for VHL treatment because it can stop or even turn back the growth of tumours. They also explained that belzutifan helps avoid surgeries, lowers the risk of metastasis and reduces the need for dialysis. The committee acknowledged the benefits offered by belzutifan. It considered that it was uncertain whether these had been fully captured in the model, given the uncertainties in the evidence and use of several assumptions in the model. The committee concluded that belzutifan could offer some additional benefits, including reduced anxiety associated with scans and fear of disability for people with VHL and their caregivers. But it noted that it had not been presented with evidence of any additional benefits specific to belzutifan that were not captured in the measurement of the QALY.

Conclusion

Recommendation

3.20 The committee's concerns about the clinical evidence and cost-effectiveness model meant that it was not confident about the results presented. It concluded that it would like the uncertainties to be addressed and that belzutifan could not be recommended.

4 Evaluation committee members and NICE project team

Evaluation committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee B.

Committee members are asked to declare any interests in the technology being evaluated. If it is considered there is a conflict of interest, the member is excluded from participating further in that evaluation.

The minutes of each evaluation committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Chair

Charles Crawley

Chair, technology appraisal committee B

NICE project team

Each evaluation is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the evaluation), a technical adviser and a project manager.

Harsimran Sarpal

Technical lead

Adam Brooke

Technical adviser

Vonda Murray

Project manager

ISBN: [to be added at publication]