NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Draft guidance consultation

Teprotumumab for treating thyroid eye disease

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using teprotumumab 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: June 2025

Note that this document is not NICE's final guidance on teprotumumab. 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- After considering these comments, the committee will prepare the final draft guidance.
- Subject to any appeal by stakeholders, the final draft guidance may be used as the basis for NICE's guidance on using teprotumumab in the NHS in England.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: Wednesday 23rd July
- Second evaluation committee meeting: TBC
- Details of the evaluation committee are given in section 4

1 Recommendations

1.1 Teprotumumab should not be used to treat moderate to severe thyroid eye disease in adults.

1.2 This recommendation is not intended to affect treatment with teprotumumab 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 healthcare professional consider it appropriate to stop.

What this means in practice

Teprotumumab is not required to be funded in the NHS in England to treat moderate to severe thyroid eye disease. It should not be used routinely in the NHS in England.

This is because there is not enough evidence to determine whether teprotumumab offers benefit in this population.

Why the committee made these recommendations

For this evaluation, the company asked for teprotumumab to be considered only for people with active moderate to severe thyroid eye disease. This does not include everyone who it is licensed for. The committee also concluded that it would only consider teprotumumab as a first-line treatment because this is what the evidence relates to.

Usual treatment for active moderate to severe thyroid eye disease at first line is methylprednisolone with or without mycophenolate.

Clinical trial evidence shows that, compared with placebo, teprotumumab reduces proptosis (bulging eyes) more and reduces diplopia (double vision) in more people.

Teprotumumab has not been directly compared in a clinical trial with methylprednisolone with or without mycophenolate. Indirect comparisons suggest that it is likely to work better than these at improving diplopia and proptosis. But the results are uncertain.

There are also uncertainties in the economic model. This is because the model does not accurately capture the disease course and how it changes with treatment, and because of some simplifying assumptions. Also, the cost-effectiveness estimates are unknown because of the uncertainty in the clinical-effectiveness evidence and modelling. So, teprotumumab should not be used.

2 Information about teprotumumab

Marketing authorisation indication

2.1 Teprotumumab (Tepezza, Amgen) is indicated 'in adults for the treatment of moderate to severe Thyroid Eye Disease (TED)'.

Dosage in the marketing authorisation

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

Price

2.3 The list price of teprotumumab 500-mg powder for concentrate for solution for infusion is commercial in confidence so cannot be reported here. The company has a commercial arrangement, which would have applied if teprotumumab had been recommended.

Carbon Reduction Plan

2.4 Information on the Carbon Reduction Plan for UK carbon emissions for Amgen will be included here when guidance is published.

3 Committee discussion

The evaluation committee considered evidence submitted by Amgen, 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.

The condition

Thyroid eye disease

3.1 Thyroid eye disease (TED) is an autoimmune condition characterised by inflammation and swelling of the muscles and tissues around the eyes. It can lead to diplopia (double vision) and proptosis (bulging eyes). Other symptoms of TED can include pain, swelling, dry eyes, redness of eyelids and ocular surface, corneal ulceration and optic nerve dysfunction. The clinical experts said that, according to the 2021 European Group on Graves' Orbitopathy (EUGOGO) guidelines, TED can be classified as mild, moderate to severe or very severe. This is based on soft tissue involvement, ocular motility, proptosis and visual function. When TED is very severe it can be sight threatening. People with TED have an active (acute) stage typically lasting up to 3 years, followed by an inactive (chronic) stage in which there are residual symptoms. The company said that, in NHS clinical practice, disease activity is also defined according to the clinical activity score (CAS). This is a scoring system composed of 7 items in which active TED is defined by a score of 3 or more. The clinical experts confirmed this is also the case. Both the company and clinical experts noted that the CAS relies on signs of inflammation and redness. This may be more difficult to detect in people with black or brown skin tones (see section 3.19). The patient experts highlighted that TED can have a substantial physical, emotional and financial impact. They explained that it can cause constant pain around the eyes. They added that proptosis can cause people to feel embarrassed about how they look, leading to reduced social confidence and activity. The patient experts further explained that diplopia impacts their ability to drive, work in bright areas and perform tasks on computers. This can limit a person's mobility

and independence and, in some cases, can lead to them losing their job or having to stop work. The committee understood that TED has significant effect on people's quality of life.

Clinical management

Treatment options for TED in the NHS

3.2 The company explained that it had received clinical opinion advising that most healthcare professionals in England follow the 2021 EUGOGO guidelines for treating active moderate to severe TED. These guidelines note the importance of starting treatment early to shorten the active phase and improve symptoms. The recommended first-line treatment for most people is intravenous methylprednisolone with mycophenolate. If symptoms are severe, a higher cumulative dose of methylprednisolone is recommended. Second-line treatment is recommended if TED does not respond to treatment or the person has signs of ophthalmic deterioration. Second-line treatments can include a second course of methylprednisolone, prednisolone with either cyclosporine or azathioprine, and orbital radiotherapy with oral or intravenous corticosteroids. The EUGOGO guidelines also recommend rituximab and tocilizumab as potential second-line treatments. But the clinical experts explained that these treatments are only available through individual funding requests, which can create barriers to using these treatments. In the inactive phase, rehabilitative surgery (orbital decompression, squint and eyelid surgery) is sometimes recommended. The aim is to attempt to restore the pre-TED appearance and function of the eyes. The clinical and patient experts explained that despite the available treatments, there is a significant unmet need in terms of efficacy and safety. They explained that current treatments aim to relieve pain, redness and swelling. But they have a limited impact on reversing proptosis and, in many people, do not improve diplopia. Also, they have significant side effects. The clinical experts added that many people with TED have to live with a disfiguring and disabling condition for years before being able to access rehabilitative surgery (when their condition is inactive). Then, even after surgery,

symptoms may persist. The committee understood that there is an unmet need for more effective treatments for people with moderate to severe TED.

Target population

3.3 Teprotumumab has a marketing authorisation as a treatment for moderate to severe TED. In its submission, the company targeted teprotumumab for a narrower population, that is people with active TED. The company said that most of the available clinical evidence is in people with active TED. It also said that the available evidence suggests that teprotumumab provides the most clinical benefit for people with active TED. The clinical experts explained that teprotumumab has been shown to be effective in both active and inactive TED but appears most effective for people with active TED. The EAG thought that restricting the population to people with active TED seemed reasonable. The committee agreed that the company's target population was appropriate.

Company's proposed positioning and comparators

3.4 The final scope issued by NICE stated that the comparator should be established clinical management without teprotumumab. This may include:

- corticosteroids (prednisolone, methylprednisolone, triamcinolone)
- immunosuppressive agents (mycophenolate, rituximab, tocilizumab, ciclosporin)
- orbital radiotherapy
- surgical interventions (eyelid surgery, orbital decompression, strabismus surgery).

The company proposed that teprotumumab should be considered as an option in the NHS for adults with active moderate to severe TED as an alternative to:

- intravenous corticosteroids (with or without immunosuppressants) at first line, and
- ciclosporin, azathioprine or orbital radiotherapy for people who have had an inadequate response to, or a deterioration in symptoms after an initial response to, intravenous corticosteroids with or without mycophenolate at second line.

At the committee meeting, the clinical experts explained that teprotumumab could be used as a first- or second-line treatment. They added that they would expect teprotumumab to be more effective at first line. But some people may prefer to have it at second line because of concerns over potential hearing loss. They also noted that the 2021 EUGOGO guidelines recommend teprotumumab as a second-line treatment. NHS England explained that it would expect teprotumumab to be used as a second-line treatment.

The company explained that methylprednisolone with or without mycophenolate should be the main comparator for teprotumumab at both first and second line. It added that almost everyone has methylprednisolone with or without mycophenolate at first line, and that a large proportion of people who need a second-line treatment also have it again. The clinical experts agreed with this. The company also explained that triamcinolone, which is injected locally around the eyes, is used only when systemic corticosteroids are contraindicated. It added that it is not widely used because of adverse events and because healthcare professionals prefer using oral corticosteroids. It also added that rituximab and tocilizumab are rarely used because they need an individual funding request in most treatment centres. The company also explained that surgical interventions are usually reserved for sight-threatening TED. The company noted that this population was not included in the licensed indication for teprotumumab.

In its submission, the company provided a scenario analysis that made

an assumption about the clinical efficacy of second-line treatments. It assumed that each of the second-line treatments plus methylprednisolone with or without mycophenolate were the same as methylprednisolone with or without mycophenolate at first line. This was because of the lack of evidence on treatments such as azathioprine, ciclosporin and orbital radiotherapy at second line. The EAG agreed that this assumption was likely to be conservative. But the clinical experts noted that people moving to the second-line treatment would usually have more severe TED than people having first-line treatment. So, it was difficult to know whether the company's assumption of equivalent efficacy would hold.

The committee noted the different opinions on the anticipated positioning of teprotumumab in the treatment pathway. But it noted that:

- teprotumumab was used at first line in the company's trials
- the matching-adjusted indirect comparison (MAIC) done by the company concerned only the comparators at first line (see section 3.7)
- there was uncertainty about whether the company's assumption of equivalent efficacy at first and second lines would hold.

The committee concluded that it would only consider teprotumumab as a first-line treatment, and that methylprednisolone with or without mycophenolate was the most appropriate comparator at first line. This was because it was presented only with the evidence comparing teprotumumab with methylprednisolone with or without mycophenolate at first line. It was also because of the uncertainties in the equivalent efficacy at second line.

Clinical effectiveness

NCT01868997 and OPTIC

3.5 The clinical-effectiveness evidence for teprotumumab was mainly from a phase 2 trial (NCT01868997; n=87) and a phase 3 trial (OPTIC; n=83).

Both were multicentre double-blinded randomised placebo-controlled trials comparing teprotumumab with placebo in people with active moderate to severe TED. They consisted of a 4-week screening phase, a 24-week double-blinded treatment phase and a 48-week follow-up phase. In NCT01868997, the primary outcome was overall response defined as a reduction in CAS by 2 or more points and a greater than 2 mm reduction in proptosis. The primary outcome in OPTIC was proptosis response, defined as a reduction in proptosis of 2 mm or more from baseline in the study eye without a corresponding increase of 2 mm or more in the fellow eye.

Evidence from NCT01868997 showed that, at week 24, statistically significantly more people in the teprotumumab arm had had an overall response than people in the placebo arm (69% compared with 20%; adjusted odds ratio 8.86; $p<0.001$). OPTIC also reported overall response as a secondary outcome. It found that at 24 weeks statistically significantly more people had an overall response in the teprotumumab arm than in the placebo arm (78% compared with 7%).

Evidence from OPTIC showed that at 24 weeks statistically significantly more people in the teprotumumab arm had a proptosis response than in the placebo arm (83% compared with 10%; $p<0.001$), In both trials, people on teprotumumab had a statistically significantly greater reduction in proptosis from baseline compared with people on placebo. In NCT01868997, the mean reduction in proptosis from baseline to week 24 was 2.46 mm in the teprotumumab arm compared with 0.15 mm in the placebo arm ($p<0.001$). In OPTIC, the mean reduction in proptosis from baseline to week 24 was 2.82 mm in the teprotumumab arm compared with 0.54 mm in the placebo arm ($p<0.001$).

In both trials, statistically significantly more people having teprotumumab had a diplopia response defined as an improvement by 1 or more diplopia grade at 24 weeks compared with people having placebo. In

NCT01868997, 68% (26/38) of people having teprotumumab had a diplopia response compared with 26% (10/39) of people having placebo ($p<0.001$). In OPTIC, 68% (19/28) of people having teprotumumab had a diplopia response compared with 29% (8/28) of people having placebo ($p=0.001$). The committee concluded that the trials demonstrated that, compared with placebo, teprotumumab is effective at reducing proptosis and diplopia.

Generalisability

3.6 The committee considered the generalisability of NCT01868997 and OPTIC to NHS clinical practice. The EAG highlighted that there was minimal involvement from people with TED from the UK in NCT01868997 and that OPTIC included no people from the UK. The company said that its clinical experts had advised that the baseline characteristics, demographics and previous medical histories of the people in the trials were broadly representative of people with TED who would be eligible for teprotumumab in NHS clinical practice. But the EAG thought that the company had not provided any evidence to support this. The committee noted that, in NCT01868997, people were excluded if they had had:

- oral or intravenous corticosteroids or any other immunosuppressive treatment for any indication in the previous 3 months
- orbital irradiation
- rituximab.

In OPTIC, people were excluded if they had had:

- corticosteroids
- orbital irradiation
- surgery.

The committee also noted that both trials included people with a clinical diagnosis of Graves' disease associated with TED and with a CAS of 4 or more. In addition, for people on teprotumumab, the mean duration

of eye symptoms or signs was 4.7 months in NCT01868997, and was 6.2 months in OPTIC. The committee recalled that, in NHS clinical practice, active TED is defined as a CAS of 3 or more (see section 3.1). The clinical experts also explained that the people in their specialist clinics, where teprotumumab would be prescribed, will usually have had TED for longer than 6 months because of NHS waiting lists and possible earlier misdiagnosis. They further explained that they would expect teprotumumab to be more effective in people who start treatment soon after symptoms appear. The committee noted that the populations in the trials may not be generalisable to NHS clinical practice. It asked the company to provide evidence to support its statement that the results from the trial populations (that is, with a CAS of 4 or more) was generalisable to people in NHS clinical practice (that is, with a CAS of 3 or more).

Indirect treatment comparison

3.7 There was no evidence from a head-to-head comparison between teprotumumab and relevant comparators. So, the company did an unanchored MAIC to determine the relative effectiveness of teprotumumab compared with methylprednisolone with or without mycophenolate. The MAIC used pooled individual patient-level data from NCT01868997 and OPTIC, and included 8 studies on methylprednisolone with or without mycophenolate. The results suggested that teprotumumab is associated with a statistically significantly higher mean change from baseline proptosis measurement when compared with methylprednisolone with or without mycophenolate. The results also suggested that people on teprotumumab are more likely to have a diplopia response than people on methylprednisolone without mycophenolate. The company considers the exact results of the MAIC to be confidential so they cannot be reported here. The committee understood the analyses suggest that teprotumumab may be effective at reducing proptosis compared with methylprednisolone with or without mycophenolate and at reducing diplopia compared with methylprednisolone alone. But the committee noted the substantial

uncertainties associated with the MAIC. These included the statistical methods employed, the population to which NCT01868997 and OPTIC IPD was matched, and the adjustment of some important prognostic variables in the analysis (see sections 3.8 and 3.9).

Adjustment for variables in the indirect treatment comparison

3.8 The covariates adjusted for in the company's MAIC comparing teprotumumab with methylprednisolone with mycophenolate were the proportion of smokers, baseline diplopia and baseline proptosis. The same covariate plus radioiodine therapy was considered in the MAIC comparing teprotumumab with methylprednisolone without mycophenolate. The company explained that treatment effect modifiers and prognostic variables were identified based on a published teprotumumab MAIC and supported by expert opinions. The company also said that the scenario analysis it had done confirmed that excluding some important covariates would result in limited residual confounding bias. But the EAG raised concerns that some important prognostic variables were not adjusted for in the MAICs. It stated that diabetes and thyroid function were not adjusted for because of limited data availability. Also, duration of TED symptoms was not adjusted for given the lack of overlaps in duration of TED symptoms across studies included in the MAICs. The EAG thought that not adjusting for these prognostic variables may have introduced uncertainties into the results of the MAIC, which were then used to generate some of the transition probabilities used in the company's model (see section 3.11). The committee understood that, given the limited availability of data, diabetes and thyroid function could not be accounted for in the MAIC. It agreed that this introduced an unknown amount of bias into the MAICs. But it noted that the lack of overlap in the duration of TED symptoms indicated heterogeneity in populations across the studies included in the MAICs. The committee recalled the statement from the clinical experts that they would expect teprotumumab to be more effective in people who start treatment soon after symptoms appear. So the committee considered it highly likely that

duration of TED symptoms is a treatment effect modifier. The committee also noted that it was unclear whether the mean durations of TED symptoms in the 2 teprotumumab trials were longer or shorter than those of the 8 studies on methylprednisolone with or without mycophenolate. It also noted that the company used pooled individual patient-level data from the 2 teprotumumab trials but aggregated data from the methylprednisolone with or without mycophenolate trials in the MAICs. But the statistical methods employed to enable this comparison and the methods used to pool the aggregated methylprednisolone trials were not clearly reported. So it is unclear which population the teprotumumab trials have been weighted to. The committee noted the substantial uncertainties in the company's MAICs, including the lack of adjustment of important covariates and insufficient reporting of the statistical methods employed. It took these into account in its decision making.

Alternative indirect treatment comparison methods and requested analyses

3.9 The EAG noted that NICE Decision Support Unit technical support document 18 lists MAICs and simulated treatment comparisons (STCs) as methods that can be used to do unanchored indirect comparisons. The company only did a MAIC (see section 3.7). The company stated that NICE had not expressed a preference for a particular methodology and there was little evidence in the literature to suggest either methodology was superior. The company further stated that clinical advice to it also agreed that the results of the MAIC were clinically plausible. The EAG suggested that scenario analyses using the STC method should be done. It considered that an STC would help determine whether the results of the MAIC were robust. The committee understood that this was because the STC method would be able to ensure simulation over the full aggregate data covariate distribution, even when dealing with complex covariate structures. The committee also noted the limitations associated with the data availability (see section 3.8) would remain regardless of the adjustment method used. But because the comparator trials included in

the MAICs cover a range of different TED durations, which spans the duration of the teprotumumab trials, theoretically it is also likely that the STC might be more appropriate than the MAIC, which did not adjust for duration of TED. So, it thought that STC methodology may be better able to account for the lack of overlap in the duration of TED symptoms across studies. Considering the substantial uncertainties associated with the MAICs (see section 3.7), the committee asked the company to:

- provide additional information on how the MAICs were done, including characteristics of populations at baseline across the trials included in the MAICs, such as duration of TED symptoms and other important variables that may confound treatment effect
- clearly report on the statistical methods used and how the aggregate data were pooled that enabled the comparison between the pooled individual patient-level data from the 2 teprotumumab trials and the aggregated data at trial level from the 8 comparator trials
- provide analysis such as an STC with and without adjustment for duration of TED symptoms and exploring both linear and non-linear relationships to determine the robustness of the MAICs
- consider, of those trials included in the MAICs, if there are any individual trials for the comparator where the baseline characteristics are more similar to those in the teprotumumab trial, and conduct a naive comparison between these.

Economic model

Company's modelling approach

3.10 The company did a Markov model to estimate the cost effectiveness of teprotumumab compared with methylprednisolone with or without mycophenolate mofetil as first-line treatment. The model had a cycle length of 6 weeks, a lifetime time horizon and 9 mutually exclusive health states. These included 6 active TED health states, surgery and post-surgery to reflect the treatment pathway for some people, and death. People entered the model in 1 of the 6 active TED health states defined

by proptosis and diplopia severity. At each model cycle people could transition to any of the active disease health states or remain in their current health state. A proportion of people transitioned to the surgery health state at 3 years where they remained for 3.5 years before transitioning to a post-surgery health state. Most people in the model remained in 1 of the 6 active health states without transitioning to the surgery and post-surgery health states for the lifetime horizon, or until death. People could enter the death health state at any time. The company applied general population mortality data for all treatment arms. The company stated that TED is not directly associated with increased mortality, but people with TED are known to have a higher risk of suicide. It considered it possible that teprotumumab may improve quality of life and reduce the risk of suicide. The company said that this potential benefit was not captured in the economic analysis. It also said that the model did not formally model the transition from active to chronic TED.

At the committee meeting, the company explained that the post-surgery health state could be considered as a proxy for inactive TED. The committee understood that, according to 2021 EUGOGO guidelines, for people with active moderate to severe TED, the goal of first-line treatment is to shorten the active phase of TED and improve eye signs and symptoms. Also, treatment effect is usually better if TED is treated early. For people whose condition does not respond to first-line treatment, second-line treatment using different medications, alone or in combination, may be considered. If the condition remains unresponsive after that, surgery may be offered. The committee noted that the company did not model second-line treatments in the model, and not everyone with inactive TED would have surgery. The committee heard that surgery might be a proxy for TED stabilisation in the company's model, which would usually improve people's appearance and potentially their quality of life. One of the clinical experts confirmed that people with TED may still think themselves disfigured even after the surgery, but usually quality of life improved. But the committee noted that this potential improvement in

quality of life may not have been reflected in the model because the utility values for active and post-surgery health states were similar (see section 3.15). The committee did not consider that the company's model captured the natural history of TED, nor the treatment pathway for the condition or the role of surgery in it. It would like to see a revised model that reflects the natural history and prognosis of TED and the treatment pathway with the transition to inactive disease explicitly considered.

Transition probabilities

3.11 The company used observed data from NCT01868997 and OPTIC to derive the transition probabilities up to week 24 for people who had teprotumumab in the model. The company said that this approach made use of all available data and produced outcomes that most accurately reflected the clinical trial data. The EAG stated that it had concerns with the approach the company had taken to model teprotumumab transitions. It noted that the company's approach generated zero percent transition probabilities for many transitions, which implied that these transitions could not happen. It also noted that there were large variations in transition probabilities between model cycles. The EAG noted that the transition probabilities for methylprednisolone with or without mycophenolate were estimated based on the teprotumumab transitions and the results of the MAIC. It had concerns about the teprotumumab transitions and thought the MAIC results were associated with uncertainty. The EAG noted that, to derive transition probabilities for methylprednisolone with or without mycophenolate, the company made various simplifying assumptions. The EAG considered that this introduced additional unquantifiable uncertainty. The EAG proposed alternative approaches to modelling transition probabilities. The company stated that there are issues with alternative approaches (such as producing a regression equation), including that they:

- are more complex
- need additional assumptions

- would still be associated with uncertainty because of the small sample size.

The EAG suggested that using a cycle length longer than 6 weeks would reduce the small sample size issue. It also suggested that modelling transition directly up to week 24 would have avoided the issue of zero percent transition probabilities. The committee concluded there was uncertainty in the modelling of transition probabilities. It would like the company to explore the impact of cycle lengths on zero percent transition probabilities in the model by varying cycle length, including directly up to week 24.

Modelling long-term treatment effect

3.12 In its model, the company assumed that, after 24 weeks when treatment stops, people remained in the same health state unless they transitioned to the surgery or death health states. It said that this assumption was supported by the available evidence. The company thought it plausible that, when TED responds to teprotumumab it would stabilise in most people. The company explained that, as people move into the inactive phase of TED, usually between 1 and 3 years after diagnosis, their symptoms stabilise. The EAG disagreed with the company's assumption. It noted that 72-week follow-up data from NCT01868997 and OPTIC suggested that, between 24 and 72 weeks, there was a reduction in proptosis and in diplopia response in people having teprotumumab. This reduction in response was greater if people who did not complete the follow-up period to 72 weeks were assumed to be 'non-responders' and were included in the calculation. (The data is considered confidential so cannot be reported here.)

At the committee meeting, the company acknowledged that evidence from NCT01868997 and OPTIC suggested that there was some reduction in response at 72 weeks. But it added that the evidence from the extension of OPTIC showed improved proptosis response. The company could not

explain why the assumption of no change in treatment effect beyond 24 weeks was made when questioned by the committee. It said that, given the available data, deriving long-term transitional probabilities for both teprotumumab and methylprednisolone with or without mycophenolate would introduce additional uncertainty. It also noted that the MAIC only reported outcomes for methylprednisolone at 12 weeks and for methylprednisolone with mycophenolate at 24 weeks. The EAG noted that the company had not provided the 72-week follow-up data in a format that could be directly incorporated into the current modelling. But, it thought that it was essential to explore the impact of treatment effect waning in the longer term.

The EAG provided a series of scenario analyses ranging from assuming people remained in the same health states after 3 years to modelling a 30% reduction in treatment effect in the longer term. It noted that the cost-effectiveness estimates were highly sensitive to health-state occupancy at 24 weeks (see section 3.11) and the assumption of treatment effect waning. It also noted that the 30% reduction in treatment effect waning may provide an upper bound for the cost-effectiveness estimates. The committee noted the comments from NHS England and the clinical experts that there was evidence that, for some people, symptoms return after stopping teprotumumab. It was also aware that, in a recent study by Ugradar et al. (2024), evidence showed that about 24% of people who completed a course of teprotumumab needed retreatment. So, it considered that the EAG's upper bound of 30% reduction in treatment effect may still not have captured the magnitude of loss in treatment effect in the longer term. Considering the evidence available and the uncertainties, the committee concluded that treatment waning should have been accounted for in the model. It also asked the company to provide evidence on teprotumumab's treatment effect in the longer term. In particular, it would like to see evidence at 72-week follow up for the intention-to-treat population in trials so that the trajectory of

teprotumumab's treatment effect in the longer term beyond 72 weeks can be assessed and incorporated in the economic model.

Time horizon

3.13 The committee noted that the company's base case adopted a lifetime time horizon. The committee noted that the cumulative effect of assuming a lifetime time horizon and assuming that, after 24 weeks, people remained in the same health state had a substantial impact on the cost-effectiveness estimates (see section 3.12). The company said that it chose to adopt a lifetime time horizon because this is needed by the NICE reference case. At the committee meeting, the NICE technical team explained that the reference case does not mandate a lifetime time horizon. Instead, it states that the time horizon should be long enough to reflect all important differences in costs or outcomes between the technologies being compared. The committee recalled that section 4.2.25 of NICE's health technology evaluations manual states that 'a time horizon shorter than a patient's lifetime could be justified if there is no differential mortality effect between technologies and the differences in costs and clinical outcomes relate to a relatively short period'. The committee noted that:

- people stopped treatment at 24 weeks in the company's model
- the company assumed no change in the treatment effect in the longer term after stopping treatment
- in the company's base case, it was assumed that there was no differential mortality because general population mortality was applied for all treatment arms.

Considering the entirety of the evidence presented and the uncertainties in the company's assumptions, the committee asked the company to provide scenario analyses exploring alternative time horizons.

Probabilistic sensitivity analysis

3.14 The EAG noted some concerns with how the probabilistic sensitivity analysis was implemented in the model. The first was that the probabilistic analysis assumed that the zero percent transition probabilities were fixed in each simulation. The EAG explained that by fixing all zero percent transition probabilities the company had failed to quantify the associated uncertainty. The company explained that using an informed prior distribution obtained through formal expert elicitation would help to quantify the uncertainty. But the company said that, given time constraints, it was not feasible to do a formal expert elicitation. The company also said that using an uninformed prior would have a disproportionately large impact on the transition probabilities. The EAG disagreed. It explained that adding 0.1 person to each cell had a small impact on the point estimates, but generated transition probabilities that ranged from 0% to 5% in each simulation. Also, it noted that to run the probabilistic sensitivity analysis the model needed individual patient-level data from the teprotumumab trial. The company said that this data could not be shared for confidentiality reasons. Instead, the company provided synthetic data. The EAG said that the synthetic data was provided in an alternative format and it was too late for it to be incorporated into its analyses. The EAG explained that this meant it was unable to do probabilistic sensitivity analysis or explore the impact of its non-informative prior. The committee noted the importance of a probabilistic sensitivity analysis and its role in exploring uncertainty in parameter values. The committee asked that the company work with the EAG to ensure it can independently generate accurate probabilistic results for its revised model. It also asked that the company provide additional scenario analysis in which 0.1 people are added to each cell.

Utility values

Source of utility values

3.15 Health-related quality-of-life data was collected in OPTIC. But it was collected using the Graves' ophthalmopathy-specific quality-of-life (GO-QOL) questionnaire, which is a disease-specific measure of health-

related quality of life for TED. The company explained that there was no mapping algorithm available to transform GO-QOL scores to EQ-5D. It added that, even if a mapping algorithm were available, it would not overcome the known limitations of EQ-5D when applied to eye conditions. The company noted that utility values are available from several NICE appraisals for other eye-related diseases that affect visual acuity. But the company did not consider these utilities to be relevant because they focused on visual acuity alone and did not capture the disfigurement, inflammation and pain associated with TED, which significantly impact health-related quality of life. So, the company did a time trade-off vignette study to get utility values for the health states in the model. The study recruited 111 people who were described as a demographically representative sampling of the US general population. The company explained that the vignettes were developed using the baseline clinical features of people in the teprotumumab trials and patient interviews. It added that the generated vignettes and the utilities were validated by clinical experts. Published utility values were not available for the surgery and post-surgery health states. So, utilities for the surgery health state were assumed to be the average of the 4 lowest utility values and for the post-surgery health state were assumed to be the average of the 4 highest utility values for the active disease health states. The EAG noted that the vignette study produced relatively low utility values. The EAG stated that it was concerned that the trade-off task may have been a cognitive challenge for many respondents. It also stated that the length of the health-state descriptions may have meant that respondents did not notice the difference between vignettes.

The committee agreed with the EAG that the available vignette was long and the utilities appeared to be low. But, the committee noted that the model assumed there is no differential mortality effect between teprotumumab and the comparators. So, the committee was more concerned about the relative difference between the utility values for different health states than it was about the low overall utility values. The

EAG explained that the description of vignettes had captured most of the relevant domains of health-related quality of life related to TED, including sleeping, pain and appearance. But the committee noted that the health-state vignettes had not captured domains of the EQ-5D that were unaffected by TED. The committee noted that these domains may also affect the quality of life of people with TED and consequently the validity of the utility values. The committee also noted that the company's model included disutility values for grade 3 or above adverse events with an incidence of 5% or more in the trial. No adverse events met this criterion for teprotumumab. The committee was aware that there was no systematic screening or monitoring for hearing loss in the trials. But real-world evidence (Sears et al. 2022) has shown that that up to 81.5% of people having teprotumumab have hearing-related issues. The committee recalled statements from patient experts that people would likely accept the risk of hearing issues to potentially prevent vision loss and reduce other symptoms. But the committee considered that hearing loss would likely be associated with substantial disutility, which was not accounted for in the modelling. It concluded that there were uncertainties in the company's utility values derived from the vignette study and that it would take this into account in its decision making. The committee asked the company to provide additional analyses accounting for disutilities associated with hearing loss and other adverse events experienced by people having treatment in NHS clinical practice.

Costs

Resource use

3.16 The committee recalled its discussion on the significance of hearing loss associated with teprotumumab. It noted that the summary of product characteristics for teprotumumab recommends that hearing is assessed before, during and after treatment with teprotumumab. The clinical experts explained that their experience of prescribing teprotumumab is limited. So, they could not accurately estimate the number of hearing appointments someone having teprotumumab would need. The committee noted that

the company had included the cost of 3 hearing appointments over the first 210 weeks in the model. The committee thought that this might have underestimated the number of hearing appointments a person having teprotumumab would need. This was because of the reported proportion of people having teprotumumab who experience hearing loss (see section 3.15). The committee thought it likely that treating these hearing issues would be associated with resource use and costs that were not included in the company's model. The committee concluded that costs associated with treating hearing loss and any adverse events should be accounted for in the model. It asked the company to include resource use and costs associated with hearing appointments, treating hearing issues and other grade 1 or 2 adverse events that might impact on cost-effectiveness in its model.

Cost-effectiveness estimates

Uncertainties in the evidence and modelling, and requested analysis

3.17 The committee noted the substantial uncertainties in the evidence base and in the company's modelling assumptions. It also requested additional information, evidence or further analysis, specifically:

- the company's assumption that the clinical efficacy of each second-line treatment (such as azathioprine, ciclosporin and orbital radiotherapy, plus methylprednisolone with or without mycophenolate) is the same as methylprednisolone with or without mycophenolate at first line (see section 3.4)
- the company's statement that the populations in the company's trials are generalisable to populations in NHS clinical practice; it asked the company to provide evidence to support this (see section 3.6)
- the MAICs, including the statistical methods employed, the population to which the teprotumumab data is being matched, and the lack of adjustment of some important prognostic variables in the analysis; it asked the company to:
  - provide additional information on how the MAICs were done

  - provide analysis, such as an STC with and without adjustment for duration of TED symptoms and exploring both linear and non-linear relationships, to determine the robustness of the MAICs
  - consider, of the studies included in the MAICs, if there are any individual trials for the comparator where the baseline characteristics are more similar to those in the teprotumumab trial, and conduct a naive comparison between these (see sections 3.7 to 3.9)
- the company's modelling approach; it asked the company to:
  - explore a revised model reflecting the natural history and prognosis of TED and the treatment pathway, with the transition to inactive disease explicitly considered (see section 3.10)
  - provide scenario analyses exploring alternative time horizons (see section 3.13)
  - work with the EAG to ensure that it can independently generate accurate probabilistic results for the company's revised model
  - provide an additional scenario analysis in which 0.1 people are added to each cell (see section 3.14)
- the modelling of transition probabilities; it asked the company to explore the impact of cycle lengths on zero percent transition probabilities in the model by varying cycle length including directly up to week 24 (see section 3.11)
- teprotumumab's treatment effect in the longer term and modelling of it; it asked the company to provide any evidence on teprotumumab's treatment effect in the longer term, particularly evidence at 72 weeks' follow up for the intention-to-treat population in trials, so that the trajectory of teprotumumab's treatment effect in the longer term beyond 72 weeks could be assessed and incorporated in the economic model (see section 3.12)
- the utility values derived from the vignette study; it asked the company to provide a scenario analysis that more accurately reflects the disutilities associated with adverse events such as hearing loss experienced by people having treatment in NHS clinical practice (see section 3.15)

- costs associated with hearing loss and other potential adverse events not accounted for in the model; it asked the company to include resource use and costs associated with hearing appointments, treating hearing issues and other grade 1 or 2 adverse events that might impact on cost-effectiveness in its model (see section 3.16).

Acceptable ICER

3.18 All the incremental cost-effectiveness ratios (ICERs) presented to the committee were above the range normally considered cost effective. Also, the committee could not identify a plausible ICER because of the uncertainties in the clinical-effectiveness evidence and the modelling, and asked for further analyses. The committee could also not determine a preferred ICER threshold for teprotumumab for treating TED.

Other factors

Equality

3.19 TED is more prevalent in women, trans men and non-binary people born with female reproductive organs. The average age at onset is between 40 and 60 years. The company and NHS England said that women, trans men and non-binary people born with female reproductive organs of working age are disproportionately affected by TED. The clinical experts commented that a person's socioeconomic status may affect their ability to access treatments such a teprotumumab that need multiple hospital visits.

The company and the clinical experts also noted that the approaches used to define disease activity may exclude people from being able to access teprotumumab specifically:

- the CAS, which is used to define active TED, relies on signs of inflammation and redness that are more difficult to detect in people with black or brown skin

- proptosis measurements, which vary by ethnic background because of anatomical differences.

The committee understood its obligations in relation to the Equality Act 2010. It was also aware that aiming to reduce health inequalities is one of the principles guiding the development of NICE recommendations. Gender and age are protected characteristics under the Equality Act 2010. But its recommendations do not restrict access to treatment for some people over others. The committee agreed that its recommendation would not have a different impact on people protected by the equality legislation than on the wider population. Race is also a protected characteristic under the Equality Act 2010. The committee noted the potential equality issue related to how the activity and severity of TED is defined and measured in practice. The committee thought that this might be more relevant if a positive recommendation were made. The committee agreed that there were no other equality issues.

Uncaptured benefits

3.20 The committee noted that teprotumumab is the first disease-modifying treatment and the only treatment specifically approved for treating TED. It recalled that the clinical experts thought teprotumumab to be a step change in the treatment of TED. The committee thought that teprotumumab is an innovative treatment. The committee discussed whether there were any uncaptured benefits of teprotumumab. It thought that because of uncertainties in the modelling it could not be certain if there were any uncaptured benefits.

Conclusion

Recommendation

3.21 The committee noted that the ICERs for the company's base case and all the scenarios that were presented were above the range normally considered a cost-effective use of NHS resources. It also noted the high

uncertainty in the clinical-effectiveness evidence and modelling, and requested further information and analyses (see section 3.17). So, it could not recommend teprotumumab for treating active moderate to severe TED in adults.

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 A.

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

James Fotheringham

Vice chair, technology appraisal committee A

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, a project manager and an associate director.

Ross Wilkinson

Technical lead

Yelan Guo

Technical adviser

Jennifer Upton

Project manager

Janet Robertson

Associate director

ISBN: [to be added at publication]