NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Naltrexone–bupropion for managing overweight and obesity (ID757)

The following documents are made available to the consultees and commentators:

1. **Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)**

2. **Consultee and commentator comments on the Appraisal Consultation Document** from:
   - Orexigen Therapeutic
   - Department of Health (no comments)
   - Royal College of Pathologist
   - Royal College Physicians

3. **ERG response to company ACD comments**

4. **Comments on the Appraisal Consultation Document received through the NICE website**

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

Naltrexone–bupropion for managing overweight and obesity [ID757]

Single Technology Appraisal

Response to consultee, commentator and public comments on the Appraisal Consultation Document (ACD)

Type of stakeholder:

Consultees – Organisations that accept an invitation to participate in the appraisal including the companies, national professional organisations, national patient organisations, the Department of Health and the Welsh Government and relevant NHS organisations in England. Consultees can make a submission and participate in the consultation on the appraisal consultation document (ACD; if produced). All non-company consultees can nominate clinical experts and/or patient experts to verbally present their personal views to the Appraisal Committee. Company consultees can also nominate clinical experts. Representatives from NHS England and clinical commissioning groups invited to participate in the appraisal may also attend the Appraisal Committee as NHS commissioning experts. All consultees have the opportunity to consider an appeal against the final recommendations, or report any factual errors, within the final appraisal determination (FAD).

Clinical and patient experts and NHS commissioning experts – The Chair of the Appraisal Committee and the NICE project team select clinical experts and patient experts from nominations by consultees and commentators. They attend the Appraisal Committee meeting as individuals to answer questions to help clarify issues about the submitted evidence and to provide their views and experiences of the technology and/or condition. Before they attend the meeting, all experts must either submit a written statement (using a template) or indicate they agree with the submission made by their nominating organisation..

Commentators – Commentators can participate in the consultation on the ACD (if produced), but NICE does not ask them to make any submission for the appraisal. Non-company commentator organisations can nominate clinical experts and patient experts to verbally present their personal views to the Appraisal Committee. Commentator organisations representing relevant comparator technology companies can also nominate clinical experts. These organisations receive the FAD and have opportunity to report any factual errors. These organisations include comparator technology companies, Healthcare Improvement Scotland any relevant National Collaborating Centre (a group commissioned by NICE to develop clinical guidelines), other related research groups where appropriate (for example, the Medical Research Council and National Cancer Research Institute); other groups such as the NHS Confederation, the NHS Commercial Medicines Unit, the Scottish Medicines Consortium, the Medicines and Healthcare Products Regulatory Agency, the Department of Health, Social Services and Public Safety for Northern Ireland).

Public – Members of the public have the opportunity to comment on the ACD when it is posted on the Institute's web site 5 days after it is sent to consultees and commentators. These comments are usually presented to the appraisal committee in full, but NICE reserves the right to summarise and edit comments received during consultations, or not to publish them at all, where in the reasonable opinion of NICE, the comments are voluminous, publication would be unlawful or publication would be otherwise inappropriate.

Please note: Comments received in the course of consultations carried out by NICE are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the submissions that NICE has received, and are not endorsed by NICE, its officers or advisory committees.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
1	Company	Orexigen Therapeutics	Thank you for the opportunity to comment on the April 2017 draft Appraisal Consultation Document (ACD) for the ongoing single technology appraisal (STA) for naltrexone-bupropion (prolonged release) (NB32) for managing overweight and obesity [ID757]. We are pleased that the NICE committee acknowledged that NB32 provides an innovative option after lifestyle measures have failed, and that there is a need for new treatment options for obesity. Our response to the ACD is outlined below. **Has all of the relevant evidence been taken into account?** The ACD did not mention standard management (SM) as a relevant comparator to NB32, though it is a comparator in the NICE Final Scope. SM is defined by experts in the UK as consisting of a reduced calorie intake diet and exercise. Additionally, at the 6 April Appraisal Committee Meeting (ACM), both the clinical expert, Professor John Wilding, and the patient representative from Helping people overcome obesity problems, Sarah Le Brocq, highlighted the challenges of orlistat use in a real world setting, helping to explain why it is not frequently used in clinical practice in England. Taking this into account, Professor Wilding considered SM as a relevant comparator for NB32. Considering this, we would ask the committee to consider SM as a relevant comparator for NB32. To enable the committee to see the results of its preferred economic analysis, including their stated preferences for analysis assumptions, we have re-implemented the economic model in a more efficient framework. The model now run calculations performed directly in VBA rather than reading formulae from	Comment noted. The committee has considered clinical practice and the appropriateness of the comparators. The committee accepted that lifestyle measures alone is the main comparator. Please see sections 3.1, 3.2, 3.3 and 3.4 of the final appraisal determination (FAD). The committee agreed that the updated model was implemented correctly and is appropriate for decision making. Please see FAD section 3.11

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			spreadsheets in order to perform necessary calculations (DICE methodology). As communicated with the NICE Project Manager on 12 May 2017, the gains in model run time this has provided are necessary to allow the committee to have confidence in the robustness of economic results. We believe, based on the analyses we present, that the committee will now be able to make an informed decision on NB32 as an innovative treatment for overweight and obese patients. We stressed in our company submission that the ability of the economic analysis to capture the long-term health and healthcare cost implications of weight loss is particularly limited. The analysis uses available data to link weight loss to cardiovascular event risks, and to the onset of Type 2 diabetes for non-diabetic patients, but is blind to the costs and HRQoL benefits of weight reduction in obese and overweight patients for known risks associated with over sixty health events1, including numerous cancers2, 3, hypertension and hyperlipidaemia4, joint and spinal complaints5-8, and sleep apnoea.8 In addition, the model does not take into account for an increased risk of death upon incidence of MI or stroke. Further, while the direct influence of BMI upon risk of cardiovascular death is incorporated for the first 15 years of the time horizon, beyond this the relationship between BMI and mortality risk is not captured within the model. As such, the cost-effectiveness of NB32, which offers a significant and clinically relevant weight loss benefit, is inherently underestimated. These limitations, and implications for the results, were not included in the ACD initial assessment. We trust that the committee will take into account this relevant evidence and acknowledge the underestimation of the health and healthcare cost benefits of weight loss that is implicit in the economic analysis. **Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?** In general, the summaries of clinical- and cost-effectiveness in the draft ACD are reasonable interpretations of the evidence available	The committee's consideration of these potential underestimates has been added to the FAD, see section 3.18 Comments noted.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			to committee, and we thank the committee for this. However, there are some important inaccuracies and potentially misleading statements that we highlight throughout this response. **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 race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?** If all of the uncertainty and implications of necessary assumptions embedded in the economic analysis are considered by the committee in their decision making, then we do not believe there to be a risk of discrimination. If, alternatively, the implications of inherent underestimation of the health and health care cost benefits of weight loss in the economic anlaysis is not factored into decision making, then there is a risk that discrimination on the basis of age will effectively unfold. Our reasoning for this is that NICE committees regularly appraise treatments for incurable illnesses, in which the the best plausible incremental health benefits are captured through extrapolation of pivotal trial data. As a result, incorporating the expected long-term benefits into the analysis is reliant on data outside of pivotal RCT data, and is often extremely challenging. This in itself implies that treatments with preventative benefits, such as those observed with weight loss, will be undervalued relative to treatments for incurable, end-of-life illnesses (beyond the NICE QALY weighting for end-of-life treatments). This has demonstrably been the case in this appraisal. As preventative treatments, such as NB32, tend to be used in a younger population than treatments for incurable, end-of-life illnesses, discrimination by age is likely to occur. **Are the recommendations sound and a suitable basis for guidance to the NHS?** We do not believe that the recommendations are suitable, firstly as	Comment noted. The committee considered these potential underestimates (see FAD section 3.18) but did not consider the recommendations to be discriminatory of different ages group's because there are no differing recommendations based on age.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			they are based only on the comparison with orlistat. In our response to the ACD we present additional economic analyses that are intended to help the committee to make an informed evaluation of the value of NB32 as a treatment option for NHS patients in the second Apprasial Committee meeting (ACM) on 8 June 2017.	Comment noted. The committee has reviewed the additional evidence provided by the company. Please see sections 3.15, 3.16, 3.17 and 3.18 of the FAD.
2	Consultee	Royal College of Physicians	The RCP is grateful for the opportunity to respond to the above consultation. We have liaised with our Nutrition Committee and would like to make the following comments. • There is significant unmet need for pharmacotherapy for obesity; there is a strong patient voice that is asking for this. • The only currently available treatment (orlistat) that was used for comparison is not tolerated by a very high proportion of patients. • The valid comparator should be lifestyle intervention, not orlistat. • One possibility would be to support use in patients who do not respond to or are intolerant of orlistat. • It is difficult to get into the discussion about the economic data as it needs more work.	Comments noted. The committee accepted that lifestyle measures alone was the main comparator. Please see sections 3.1, 3.2, 3.3, 3.4 and 3.20 in the FAD for the committee's key conclusions.
3	Consultee	Royal College of Pathologists	**1. Has all of the relevant evidence been taken into account?** I am not aware of any relevant studies other than those already included. **2. Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?** The summaries appear to be appropriate. The evidence demonstrates naltrexone-bupropion to be clinically effective in comparison to placebo, and indirect comparisons suggest, on the whole, similar clinical efficacy between naltrexone-bupropion and orlistat. With regard to assessment of cost effectiveness, it is very	Comment noted. Please see sections 3.1, 3.2, 3.3, 3.4 and 3.20 in the FAD for the committee's key conclusions.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			disappointing that it has not been possible to make a sufficiently reliable assessment of cost effectiveness, as it is on this basis that the conclusion has been reached that naltrexone-bupropion cannot be recommended as an option for managing overweight and obesity within the NHS. However, the basis on which this conclusion has been reached is clearly explained. **3. Are the provisional recommendations sound and a suitable basis for guidance to the NHS?** There is a clear clinical need for novel pharmacological approaches to treatment of overweight and obesity, and naltrexone-bupropion is a novel treatment that could be well placed as part of an integrated weight management pathway. It must always be emphasised that it is important for weight management services to be available within a cohesive four tier pathway, including specialist non-surgical specialist weight management services at Tier 3 and bariatric surgery at Tier 4. Against this background, it is deeply disappointing that naltrexone-bupropion cannot currently be recommended. However, in the absence of a reliable estimate of cost effectiveness, this recommendation does appear to be appropriate.	
4	Commentator	Web comment - University of Liverpool and University Hospital Aintree	Working as a metabolic physician with a specialist interest in Obesity and type 2 diabetes as well as being an active obesity clinical researcher, the lack of therapeutic options in obesity in the UK is startling. Lifestyle intervention is undoubtedly effective but in reality the proportion in whom it is effective is very limited and of limited magnitude. We have a significant unmet need with up to 4 drugs available in countries like the US. The use of orlisatat is incredibly limited by GPs or specialists and is poorly tolerated. For this reason I believe this drug should be made available as a therapeutic option. I would also suggest comparing the response against orlistat is not valid due to the very small number of patients on it. I would suggest lifestyle is the relevant comparator. Application of appropriate stopping rules ensure this drug would be	Comments noted. The committee accepted that lifestyle measures alone was the main comparator. Please see sections 3.1, 3.2, 3.3, 3.4 and 3.20 in the FAD for the committee's key conclusions.

Comment number	Type of stakeholder	Organisation name	Stakeholder comment Please insert each new comment in a new row	NICE Response Please respond to each comment
			continued in those who derive clinical benefit and I would urge NICE to endorse this drug so that we may impact upon the obesity epidemic.	

Orexigen Response to:

National Institute for Health and Care Excellence

Appraisal Consultation Document – Naltrexone-bupropion (prolonged release) for managing overweight and obesity [ID757]

May 2017

Dear Committee Members,

Thank you for the opportunity to comment on the April 2017 draft Appraisal Consultation Document (ACD) for the ongoing single technology appraisal (STA) for naltrexone-bupropion (prolonged release) (NB32) for managing overweight and obesity [ID757]. We are pleased that the NICE committee acknowledged that NB32 provides an innovative option after lifestyle measures have failed, and that there is a need for new treatment options for obesity. Our response to the ACD is outlined below.

Has all of the relevant evidence been taken into account?

The ACD did not mention standard management (SM) as a relevant comparator to NB32, though it is a comparator in the NICE Final Scope. SM is defined by experts in the UK as consisting of a reduced calorie intake diet and exercise. Additionally, at the 6 April Appraisal Committee Meeting (ACM), both the clinical expert, Professor John Wilding, and the patient representative from *Helping people overcome obesity problems*, Sarah Le Brocq, highlighted the challenges of orlistat use in a real world setting, helping to explain why it is not frequently used in clinical practice in England. Taking this into account, Professor Wilding considered SM as a relevant comparator for NB32. Considering this, we would ask the committee to consider SM as a relevant comparator for NB32.

To enable the committee to see the results of its preferred economic analysis, including their stated preferences for analysis assumptions, we have re-implemented the economic model in a more efficient framework. The model now run calculations performed directly in VBA rather than reading formulae from spreadsheets in order to perform necessary calculations (DICE methodology). As communicated with the NICE Project Manager on 12 May 2017, the gains in model run time this has provided are necessary to allow the committee to have confidence in the robustness of economic results. We believe, based on the analyses we present, that the committee will now be able to make an informed decision on NB32 as an innovative treatment for overweight and obese patients.

We stressed in our company submission that the ability of the economic analysis to capture the long-term health and healthcare cost implications of weight loss is particularly limited. The analysis uses available data to link weight loss to cardiovascular event risks, and to the onset of Type 2 diabetes for non-diabetic patients, but is blind to the costs and HRQoL benefits of weight reduction in obese and overweight patients for known risks associated with over sixty health events[1], including numerous cancers[2, 3], hypertension and hyperlipidaemia[4], joint and spinal complaints[5-8], and sleep apnoea.[8] In addition, the model does not take into account for an increased risk of death upon incidence of MI or stroke. Further, while the direct influence of BMI upon risk of cardiovascular death is incorporated for the first 15 years of the time horizon, beyond this the relationship between BMI and mortality risk is not captured within the model. As such, the cost-effectiveness of NB32, which offers a significant and clinically relevant weight loss benefit, is inherently underestimated. These limitations, and implications for the results, were not included in the ACD initial assessment. We trust that the committee will take into account this relevant evidence and acknowledge the underestimation of the health and healthcare cost benefits of weight loss that is implicit in the economic analysis.

Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?

In general, the summaries of clinical- and cost-effectiveness in the draft ACD are reasonable interpretations of the evidence available to committee, and we thank the committee for this. However, there are some important inaccuracies and potentially misleading statements that we highlight throughout this response.

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 race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

If all of the uncertainty and implications of necessary assumptions embedded in the economic analysis are considered by the committee in their decision making, then we do not believe there to be a risk of discrimination.

If, alternatively, the implications of inherent underestimation of the health and health care cost benefits of weight loss in the economic anlaysis is not factored into decision making, then there is a risk that discrimination on the basis of age will effectively unfold. Our reasoning for this is that NICE committees regularly appraise treatments for incurable illnesses, in which the the best plausible incremental health benefits are captured through extrapolation of pivotal trial data. As a result, incorporating the *expected* long-term benefits into the analysis is reliant on data outside of pivotal RCT data, and is often extremely challenging. This in itself implies that treatments with preventative benefits, such as those observed with weight loss, will be undervalued relative to treatments for incurable, end-of-life illnesses (beyond the NICE QALY weighting for end-of-life treatments). This has demonstrably been the case in this appraisal. As preventative treatments, such as NB32, tend to be used in a younger population than treatments for incurable, end-of-life illnesses, discrimination by age is likely to occur.

Are the recommendations sound and a suitable basis for guidance to the NHS?

We do not believe that the recommendations are suitable, firstly as they are based only on the comparison with orlistat. In our response to the ACD we present additional economic analyses that are intended to help the committee to make an informed evaluation of the value of NB32 as a treatment option for NHS patients in the second Apprasial Committee meeting (ACM) on 8 June 2017.

The remainder of this response is divided into six parts. Section 1 covers clarification points concerning the decision problem, and re-itearates the insurmountable limitations of the clinical comparison to orlistat. Section 2 describes the necessary reimplementation of the model in a more efficient

platform. Section 3 sets out the committee's base case preferences as we percieve them from the draft ACD as a Revised Base Case. Section 4 includes results from the Revised Base Case and notes on interpretation. Section 5 is the reference list and Section 6 is an Appendix, containing evidence of the robustness of the revised economic analysis to first-order uncertainty.

Yours sincerely,

Hans-Joerg Fugel

VP Market Access Europe,

Orexigen Therapeutics

1 The decision problem

1.1 Standard management without pharmacological treatment is the sole relevant comparator for some patients

Section 3.16 of the ACD strongly implies that the comparison to orlistat determines the cost-effectiveness of NB32 as a use of NHS resources:

"*The committee noted that the incremental results for the comparison with orlistat suggested naltrexone–bupropion is not a cost effective use of NHS resources*"

In the context of a full incremental analysis of NB32 adjunctive treatment and its two Final Scope comparators, the strength of this statement rests on the assumptions and limitations of the NB32-orlistat clinical comparison set out in Section 1.3 of this response. However, there are a group of patients for whom orlistat treatment is not a feasible option, and for these patients, a pairwise comparison to SM is relevant to determine cost-effectiveness. This group were not defined in Clinical Guideline 189,[9] but those patients who have discontinued orlistat due to adverse reaction are an easily identifiable subset of the group for whom orlistat is not a plausible treatment option.

In the 6th April ACM, the committee heard the patient expert explain the poor uptake of orlistat in the NHS, and why there is such unmet need in overweight and obese patients who are eligible for Tier 3 services. As documented in Section 3.2 of the ACD, "*its use is limited …….orlistat causes unpleasant and socially unacceptable gastrointestinal side-effects*". From this account, recommendation of NB32 will not primarily displace orlistat, but rather provide an option for patients whose options are currently limited to SM.

1.2 A treatment sequence in which NB32 is used after orlistat failure is relevant for decision making

If the committee do consider orlistat as a comparator and are unable to conclude that NB32 is a cost-effective alternative to orlistat when pharmacological adjunct is first considered after standard management (defined in Clinical Guideline 189 as dietary, exercise and behavioural approaches that have been started and evaluated),[9] the

committee may want to consider NB32 as an alternative to SM alone in patients who fail orlistat treatment.

The key points to consider for appraisal of this group are as follows:

- Patients in the COR trial programme did not have orlistat in the 4 weeks before treatment initiation, but orlistat use prior to the 4 weeks before entering the COR trails was not documented. Considering the patients enrolled in the COR program had chronic obesity, a desire to achieve weight loss, and orlistat available through the marketplace, it is fair to assume that a proportion of them had tried orlistat and not responded in a satisfactory way
- The fact that orlistat and NB32 have different mechanisms of action implies that previous orlistat treatment is not expected to affect NB32 treatment effectiveness[8]

As such, the base case pair-wise comparison to SM is an accurate estimate of the most plausible ICER for NB32 versus SM in orlistat treatment failures.

Nevertheless, as the committee expressed a need for the model to capture retreatment (ACD Section 3.3, discussed in Section 3.4 of this response), which is now possible due to reimplementation of the model in a faster platform meaning model execution time has no implications for deterministic analyses (Section 2 of this response). Section 4 of this response presents a treatment sequencing scenario for a world in which those who discontinue orlistat due to insufficient response begin NB32 treatment compared to a world in which those who discontinue orlistat due to insufficient response begin SM.

1.3 Standard management without pharmacological treatment is the only comparator for which there is credible comparative evidence

As per the NICE final scope, we provided the most complete and robust comparison to adjunctive orlistat treatment possible in our evidence submission. However, there were challenges in making a comparison of clinical outcomes across NB32 and orlistat studies, that we have previously emphasised, including post-hoc stopping rules, trials with different designs conducted years apart, and across periods when the treatment paradigm for patients, especially those with type II diabetes, has significantly changed. These important challenges raises questions on the

robustness of the comparison between the two treatments and we therefore ask the committee to carefully consider the limitations and assumptions, and the implications this has on comparing the clinical benefits of NB32 to that of orlistat.

Section 3.8 of the ACD concludes from indirect treatment comparisons (ITCs) to orlistat that there is "*similar efficacy between naltrexone–bupropion and orlistat but that orlistat may be more effective in changing mean weight in people with type 2 diabetes*", following the ERG's conclusions on the most relevant trial datasets for comparison. Yet, none of the of 1-year outcomes from the published evidence base are reflective of clinical practice, where 12- and 16-week stopping rules are in place for orlistat and NB32, respectively. Additionally, as we heard from both Sarah Le Brocq and Professor Wilding at the first ACM, orlistat is associated with unpleasant and socially distressing gastrointestinal side-effects that have limited its usefulness in clinical practice.

The post-hoc stopping rules included as part of the license description for both orlistat and NB32 in this indication make comparison of 12-month outcomes across these treatments extremely challenging, if not impossible. Patient-level data access would have allowed identification of responders and non-responders at 12/16 weeks, and comparison of outcomes across responders at study endpoints; a comparison that would reflect discontinuation rules in clinical practice. Access was only possible for NB32 trial data, while the publicly available data for orlistat trial patients stratified by 12-week response are scant. Results from the ITC are blind to the implications of these important discontinuation rules. It is plausible that the relative benefit of NB32 versus orlistat is underestimated by the results of the ITC (and therefore the economic model). We therefore ask the committee to carefully consider the limitations and assumptions implicit in the ITC, and what these mean for inference.

Consider also the stated committee preferences for orlistat time-to-treatment-discontinuation (TTD) assumptions in the economic analysis (ACD Section 3.14). The committee, following the advice of the ERG, are keen that the TTD data for NB32 should not be applied to the orlistat arm in a way that could bias in favour of NB32. Yet, considering the account given by the committee's patient expert of the side-effects associated with orlistat, does it not seem to the committee that using NB32 TTD data as a proxy for orlistat TTD data could substantially overestimate orlistat TTD? As weight loss benefit begins to deteriorate upon discontinuation, this

would overestimate the benefit of orlistat as well as the costs, and potentially underestimate the cost-effectiveness of NB32 in this comparison.

Furthermore, as described in Section 4.10.4 of the company's submission (CS), much heterogeneity is observed between the orlistat studies included in the ITC. Although the ITC explores some of this heterogeneity through subgroup analyses, differences between studies are still present. The subgroup analysis in which studies with intensive standard management without pharmacological intervention (SM) were excluded appeared to show differences in results. It is therefore also possible that more subtle differences in SM may also impact the relative effectiveness of the different treatments. The SM regimens patients received in the orlistat studies were not consistent and varied between studies in terms of number of clinician visits, recommended diet and exercise regimes, and additional weight loss aids administered such as pedometers and food diaries. The comparison between orlistat, SM and thus the comparison of orlistat with NB32 through the ITC is subject to these limitations.

In addition to the differences in orlistat study design, the analysis populations and imputation methods for the orlistat studies were also heterogeneous, and different to those preferred by the committee for base case analysis (Section 3.14 of the draft ACD). Out of the eight orlistat studies that were used to inform the economic model, the mean % weight loss from baseline, was estimated in two studies using the mean weight of patients that remained in the study for one year and the mean weight at baseline (as described in Appendix 10 of the CS).[10, 11] The mean % weight loss estimated from these two studies is therefore likely to be more representative of a completers analysis rather than an analysis in the intention-to treat (ITT) population. In the remaining six studies, two studies did not define their analysis population[12, 13] and the remaining four studies used an ITT population with restrictions around the receipt of treatment and post baseline measurements;[14-17] the last observation carried forward imputation method was used in each of these four studies. All these contrast with the baseline observation carried forward imputation method used for NB32 trial data in the ERG base case, and this biases the ITC between orlistat and NB32 in favour of orlistat.

Based upon the lack of implementation of the stopping rule and the heterogeneity in the orlistat studies, it is highly plausible that relative benefit of NB32 versus orlistat is

underestimated by the results of the ITC (and therefore the economic model). We therefore ask the committee to carefully consider the limitations and assumptions implicit in the ITC, and what these mean for inference.

It is our opinion that, considering the limitations of available orlistat data, it is not appropriate to consider orlistat as the only comparator. SM is the only credible comparator for which there is more relevant comparative evidence supported by high-quality RCT, and access to patient-level data has allowed the decision problem for this comparison to be addressed in the economic model.

2 Model implementation

Model implementation has been an unwanted challenge of this appraisal to date, and a source of frustration for the company, the ERG and the committee. We wish to express our gratitude to the committee for allowing the company representative a chance, during the 6th April 2017 ACM, to explain the causes and implications of the long model run time. We hope that upon consideration of this response, the committee can be convinced of the reliability and usefulness of the revised cost-effectiveness model for the purposes of this technology appraisal.

Section 3.10 of the ACD describes the model implementation problems as follows:

"*The ERG had concerns about how the economic model was implemented using Discrete Integrated Condition Event (DICE) methodology in Excel, which caused extremely slow run times. The company recognised the limitations of executing the model in Excel and using DICE*"

This text is potentially misleading, with important implications for future NICE Single Technology Appraisals (STAs). Executing a discrete event simulation model in Excel® is not typically assocoiated with slow model run times, when the traditional approach of implementation in *Visual Basic for Applications*® (VBA) is taken. As described in the 6th April ACM, the decision to use the DICE approach in Excel®, rather than VBA, was based on: (i) the need to find an acceptable platform for the correct model type within a restricted timeframe; (ii) reluctance expressed within previous appraisals regarding the acceptability of VBA-based Excel® models without forewarning, reiterated by the NICE Project Manager at the Decision Problem Meeting for this appraisal; and (iii) the publication disseminating the DICE approach,[18] which illustrated the transparency benefits of the approach for an audience more comfortable reviewing data and logic in spreadsheets than in underlying code. Unfortunately the publication did not indicate the model run-time issues experienced when applying DICE in standard spreadsheet software.

To allow incorporation of the committee's preferred assumptions into an analysis that can explore first- and second-order uncertainty sufficiently, we have accepted that it was necessary to re-implement the economic model in a more efficient framework. As such, we have devoted time and resources to accurately update the submitted model, so that rather than reading formulae from spreadsheets in order to perform

necessary calculations, model run calculations are performed directly within VBA code. As a consequence, run time for one simulated patient has fallen, from a peak of over 10 minutes using the DICE framework, to a fraction of a second.

Of course, this was not a straightforward update. Re-implementation of the model in VBA code was associated with challenges in aligning calculations with the original model, due to the differences in logic between VBA code and Excel worksheet functions. Some mathematical functions such as additions and multiplications were simple to re-implement using a "copy and paste" approach. Other functions such as "IF" statements were slightly more cumbersome to re-implement: nested "IF" statements in VBA code require separation over several lines of code. The next level of complexity considers those functions that do not have a VBA counterpart – for example, a vertical lookup ("VLOOKUP") function. This type of function requires VBA code to state that it is a worksheet function by including the text "Application.WorksheetFunction." ahead of the name of the function. Finally, some DICE-specific features required complete re-structuring, as the equivalent function cannot be considered in VBA. For example, in the DICE model the next event was selected using "live" values and a "live" "MIN" function. In VBA, the selection of an event had to be coded as a loop, where each event was assessed as a candidate for the next event with the smallest of these times selected based on a conditional subroutine called to actively calculate the minimum of the derived values.

The iterative process of assessing and improving the accuracy of reimplementation involved a combination of patient-by-patient analysis of condition values at run-end and comparsions of deterministic base case results for identical patient cohorts, across pre- and post-implementation models. The version of the model underpinning the ERG report, named "ID757 Naltrexone - bupropion HE ERG_base-case 090317 LG [ACIC]", was shared with Orexigen upon request following receipt of the draft ERG report. This was felt to be a reasonable starting point for reimplementation, containing the functionality to select ERG preferences, many of which the committee are in agreement with.

Table 1 shows a comparison of mean deterministic base case results from pre- and post-reimplementation versions of the model, following the iterative process of assessing and improving the accuracy of reimplementation. The results were produced based on a sample of 1,000 patients; running a sufficient number of

patients through the model to stabilise the base case ICER was a key concern *following* reimplementation, but *to test* reimplementation, the intention was for results to match exactly, irrespective of the number of patients informing a model run. The patient profiles and random number seeds used to test the reimplementation were matched to those used to run results in the DICE model “ID757 Naltrexone - bupropion HE ERG_base-case 090317 LG [ACIC]”.

Table 1 highlights that while an exact match was not quite achieved, we reached a point where key model results averaged across 1000 patients are very similar across the DICE model and the reimplemented VBA model. The ICER for NB32 versus orlistat is extremely sensitive to outcome changes, owing to an estimated health benefit of less than 0.02 QALYs; this ICER varies by less than £210 across the two models in Table 1, or less than 1%.

Table 1 presents an alternative illustration of the consistency of the pre- and post-reimplementation results. The proportion of the 1,000 patients in the Table 1 sample whose final outcomes matched exactly across NB32, orlistat and SM model runs is presented, for six key model outcomes inclung total discounted QALYs and total discounted life years, and remains above 90% across outcomes.

We truly hope that the committee and ERG are satisfied that the reimplementation has been an exhaustive attempt to meet the committee’s needs for fair and timely decision making. This was always our intention, and we again apologise that our original implementation strategy, decided upon with the best intentions, caused obstruction to the ERG and committee.

Table 1: Comparison of pre- and post-reimplementation base case model results for 1000 simulated patients

Technologies	Total			Incremental			ICER	
	Costs	LYs	QALYs	Costs	LYs	QALYs	Versus baseline [SM]	Incremental
Incremental								
DICE model "ID757 Naltrexone - bupropion HE ERG_base-case 090317 LG [ACIC]"								
SM	£5,964	34.4406	15.1111					
ORL	£6,275	34.4760	15.1950	£311	0.0354	0.0839	£3,701	£3,701
NB32	£7,017	34.4814	15.2113	£742	0.0054	0.0162	£10,510	£45,694
Reimplementation in VBA								
SM	£5,864	34.4920	15.1339					
ORL	£6,166	34.5145	15.2160	£302	0.0226	0.0822	£3,678	£3,678
NB32	£6,908	34.5198	15.2324	£742	0.0053	0.0163	£10,600	£45,488
Pairwise								
DICE model "ID757 Naltrexone - bupropion HE ERG_base-case 090317 LG [ACIC]"								
SM	£5,964	34.4406	15.1111					
NB32	£7,017	34.4814	15.2113	£1,053	0.0408	0.1002	£10,510	
ORL	£6,275	34.4760	15.1950					
NB32	£7,017	34.4814	15.2113	£742	0.0054	0.0162	£45,694	
Reimplementation in VBA								
SM	£5,864	34.4920	15.1339					
NB32	£6,908	34.5198	15.2324	£1,044	0.0278	0.0985	£10,600	
ORL	£6,166	34.5145	15.2160					
NB32	£6,908	34.5198	15.2324	£742	0.0053	0.0163	£45,488	

Key: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; LYs, life years; NB32, naltrexone bupropion; ORL, orlistat; QALYs, quality-adjusted life years; VBA, Visual Basic for Applications.

Table 2: Consistency of patient-specific model outcomes across pre- and post-reimplementation models, over 1000 simulated patients

Model Outcome (discounted)	Patient match rate
Life years	97.8%
Quality-adjusted life years	94.6%
Drug costs	100.0%
Standard management and condition management costs	90.4%
Adverse event costs	100.0%
Death costs	99.5%

3 Capturing the committee's preferences

In Sections 3.9 to 3.16 of the draft ACD, the committee helpfully set out their preferences for model structure, execution and assumptions. We attempt to specify a Revised Base Case by interpreting and incorporating these preferences into the reimplemented economic model.

The ERG made eight model changes to inform their preferred analysis in the ERG Report. These changes were listed in slides 22 and 23 of the Committee Lead Team's Cost-Effectiveness presentation from the 6 April ACM, and are reproduced with the same wording as used in the slides, in Table 3.

Table 3: ERG Amendments to Original DICE model, for ERG base case analysis

Amendment	Amendment Description
1	Fixed error in the weight regain assumption so it is regained linearly over 3 years rather than being regained instantly after 3 years
2	Used ITT population from COR-I and COR-DM trials instead of a mITT pooled population
3	Used a relative risk instead of mean differences to extrapolate the difference between treatments in change from baseline weight from the secondary to the primary assessment
4	Calibrated the BMI natural history model to reflect baseline BMI as per the COR trials (mean BMI of 36 kg/m2)
5	Adjusted baseline age, proportion of females, smokers, people taking aspirin, anti-hypertensive medication and statins using baseline characteristics from COR trial programme, stratified for T2DM status, if applicable
6	Removed GP visit cost (52-week assessment) for people receiving standard management
7	Assumed weight regain towards baseline BMI instead of predicted BMI from the natural history model, in 3 years
8	Removed linear scaling assumption for TTD for orlistat
Key: BMI, body mass index; DM, diabetes mellitus; GP, general practitioner; ITT, intention-to-treat; mITT, modified intention-to-treat; TTD, time-to-treatment discontinuation.	

How the Revised Base Case reflects other ERG preferences is set out in the remainder of this section with primary reference to the specific concerns listed in the ACD, before the differences between the CS assumptions and Revised Base Case assumptions are summarised in Section 3.8.

3.1 Simulated patient baseline characteristics

Section 3.13 of the draft ACD reiterates the committee's conclusion that participants in the COR trial programme are similar to those likely to be seen in practice in

England, and states the committee's preference for baseline patient characteristics in the model to reflect those of patients in the NB32 trials.

This committee preference is incorporated in the Revised Base Case through ERG amendments 4 and 5 from Table 3.

3.2 Treatment effectiveness and discontinuation

Section 3.14 of the draft ACD states the committee's preferences for the ERG's assumptions on treatment effectiveness and TTD.

We accept the committee's rationale, and the Revised Base Case is therefore consistent with amendents 2 and 8 from Table 3. In addition, the model has been further updated to meet the committee's preferences, in that rather than pooled data from all four COR studies, TTD data from COR-I and COR-DM only are now used to inform TTD for the first 56 weeks of the model. As such, no COR-II or COR-BMOD data are informing the Revised Base Case, meeting the commitee's preferences. Consistent with the CS and ERG's analysis, we present results for baseline T2DM and baseline non-T2DM subgroups in Section 4 of this document.

Orexigen believe that the modified ITT analysis set used in the CS base case is the most relevant analysis set to compare weight loss between patients taking NB32 and those on SM as this is the recommended FDA approach, as stated in their draft guidance for 'developing products for weight management'.[19] However, we understand the ERG's and committees reservations of this analysis and as a result agree to use ITT within the economic analysis.

3.3 Weight regain following treatment discontinuation

Weight regain trajectory

Section 3.15 of the draft ACD states the committee's preference for a return to baseline weight over 3 years following treatment discontinuation, based on the rationale that this assumption was used by Ara et al,[20] and consistent with ERG amendment 7 from Table 3.

We were clear to stress the rationale for patients to trend towards their BMI projection following treatment discontinuation in both our CS, and our ERG Report Factual Inaccuracy Check (FAC) (Issue 9). The reason we did not go with the base

case assumption used by Ara *et al* is because it seems to reject the evidence Ara *et al* synthesised on the natural history of BMI. Consider a patient who is treated with orlistat for 12 weeks, achieves 0.5% weight loss and so discontinues at their primary assessment. Assuming a return to baseline weight in 3 years would imply nearly stable weight for 3 years from a treatment that has had no discernable effect, when the evidence from Ara et al tells us that BMI increases with age.[20] After 3 years, to avoid a spike in BMI, the ERG approach underpredicts BMI relative to the BMI trajectory data for the rest of the simulated patient's life. Effectively, lifetime BMI trajectory is changed based on 12 weeks of ineffective treatment.

Our approach maintains the assumption of weight regain over 3 years, so that over the three years post treatment discontinuation, patient BMI is trending back towards the appropriate weight for that patient's characteristics, given the BMI natural history evidence we have.

For the Revised Base Case, we do not want to incorporate an assumption that diverts from evidence, rather than aligning with evidence. Further, we do not feel that the committee and ERG would want us to, after reconsideration of the implications. As such, ERG amendment 7 from Table 3 is not included in the Revised Base Case.

Weight regain implementation

ERG amendment 1 was described to the committee as correction of an "*error in the weight regain assumption so it is regained linearly over 3 years rather than being regained instantly after 3 years*". In Issue 7 of our ERG Report FAC, we explained how this was not an error, and how the ERG "error fix" introduces error to the calculation of patient utility and time-to-event estimates.

The ERG response to Issue 7 expressed concerned about "estimated times to subsequent events" with our approach. Our approach updates patient BMI at 3 years post-regain, calculates appropriate QALYs accrued based on linear interpolation, and adjusts time-to-event estimate to reflect updated patient characteristics. The notion that not updating BMI and TTE estimates for 3 years to reflect changes over time is consequential for results is highly unlikely, particularly in comparison to the implications of the ERG's fix. We describe in Section 3.6 of this response how annual BMI updating has been incorporated into the Revised Base Case. This should allay the concerns the ERG indicate in their response to Issue 7.

As implementation of amendment 1 from Table 3 would introduce error, this amendment is not included in the Revised Base Case.

3.4 The treatment pathway

Section 3.9 of the draft ACD concerns the model type and structure, concluding that the type and structure of the model are appropriate but that "episodes of retreatment and a transition to bariatric surgery should be included in the model".

Bariatric surgery

In the UK, patients are eligible for bariatric surgery if: they have a BMI of 40 kg/m^2 or more, or 35 kg/m^2 or more with another significant disease that could be improved with treatment; and all appropriate non-surgical measures have been tried, but adequate, clinically beneficial weight loss has not been achieved or maintained.[9] In practice, as explained by the clinical expert on 6th April, service capacity means bariatric surgery is used in only a tiny fraction of the eligible population. As such, it is targeted to those with the greatest potential for benefit, such as those with type 2 diabetes, sleep apnoea and very high BMI.[21] A 2016 press release by the British Medical Journal (BMJ) reported that in 2014-15 there were 6,032 bariatric surgeries carried out by the NHS.[21] An estimated 2.6 million people are eligible for bariatric surgery in the UK.[21] From these figures, an estimated 0.232% of eligible people bariatric receive surgery every year. This estimate is slightly higher than, but reflective of, the estimate of 0.1% from the 6th April ACM (Section 3.1 of the draft ACD).

Using the BMJ figures as a starting point, it has been possible to incorporate subsequent bariatric surgery into the Committee Base Case as follows:

- The outcome of bariatric surgery is either: success (weight loss); failure (no lasting weight loss); or death.
- Bariatric surgery successes are assumed to lose 24.225% of their body weight from bariatric instantly.[22] Instant weight loss is a simplifying assumption. The estimate of 24.225% is an average of 10-year weight loss outcomes reported in a *Health Technology Assessment* systematic review of bariatric surgery for obesity (Table 53).[22]
- Bariatric surgery is assumed to be fatal for 0.1% of patients treated.[22, 23]

- Bariatric surgery is assumed to fail in 12.5% of patients treated.[24] Patients who fail bariatric surgery are assumed to lose no weight but maintain current weight until death. Again, this is a simplifying assumption.
- Bariatric surgery is assumed to only be possible two years after meeting eligibility criteria, based on the expectation of a long average waiting list duration, given the elective nature of the procedure.[25]
- Bariatric surgery is assumed to cost £4,886, based on the NHS Reference Costs code FZ84Z: Stomach Bypass Procedures for Obesity, 19 years and over.[26] Follow-up costs post-surgery are assumed to comprise one GP appointment per year.[23] The cost of immediate follow-up is assumed to be captured by the NHS Reference Cost FZ84Z.
- Bariatric surgery provision is assumed to be used to treat patients with T2DM over those without T2DM, as in NHS practice it is known to be targeted to those with the greatest potential for benefit, such as those with type 2 diabetes, sleep apnoea and very high BMI.[21]

Retreatment with pharamacological adjunct

NB32 for patients who fail orlistat

As discussd in Section 1.2 of this response, the committee may wish to consider NB32 adjunct therapy as an alternative to SM alone in patients who fail orlistat treatment. Further, NB32 effectiveness data from the COR trial programme are expected to be a fair representation of outcomes for NB32 patients following orlistat failure:

- Orlistat use by patients in the COR trial programme prior to the four-week period immediately preceding investigative treatment is unknown, but can be reasonably expected to reflect clinical practice
- The different mechanisms of action of orlistat and NB32 mean that previous orlistat treatment is not expected to independently predict NB32 treatment effectiveness[8]

As such, we suggest that the base case pair-wise comparison to SM is a fair estimate of the most plausible ICER for NB32 versus SM in orlistat treatment

failures. Nevertheless, with the thought that it may be useful to the committee, we specify a scenario to test the cost effectiveness of NB32 for orlistat failures explicitly.

In this scenario, it is assumed that orlistat patients who present with <5% weight loss at their 12-week response assessment are then either treated with (i) SM or (ii) NB32 (alongside SM). The choice between these treatment options represents the decision problem for orlistat failures.

To implement this scenario, the following steps were taken. The model was run with the time horizon set to 12 weeks, and all patients were set to be non-responders at primary assessment. Otherwise assumptions and settings were consistent with the Revised Base Case. Mean orlistat results from this model run were then added to (i) SM mean model results from a full run of the Revised Base Case analysis and (ii) NB32 mean model results from a full run of the Revised Base Case analysis, separately. Incremental analysis of these two sets of results comprises the scenario analysis assessing NB32 as an alternative to SM in orlistat treatment failures; one arm represents orlistat for 12 weeks followed by NB32, and the other represents orlistat for 12 weeks followed by SM.

Retreatment with the same treatment

From the 6th April 2017 ACM and the ACD, we infer that the committee would like to consider the potential ramifications of re-challenging with the same treatment, for cost-effectiveness estimates.

Given that no data are available either on the expected frequency of retreatment, expected timing of retreatment or effectiveness of NB32 when used as a retreatment, the uncertainty that would surround the inclusion of a scenario exploring retreatment as an option would be extremely high. We therefore feel unable to provide the committee with an informative scenario analysis exploring the cost-effectiveness of NB32 when retreatment is included.

If there is no relationship between treatment effectiveness and re-treatment effectiveness, the Revised Base Case results provide a fair estimate of the cost-effectiveness of retreatment with the same treatment.

3.5 Capturing first- and second-order uncertainty

In Section 3.11 of the draft ACD, the committee conclude that far too few patient simulations were informing deterministic model results, and that deterministic results were not sufficiently reliable for decision making. The efficiency of the reimplemented model means that sufficient simulated patients can be run through the model to eliminate any worries of first-order uncertainty biasing deterministic results. Base case deterministic Revised Model Results shown in Section 4 of this response are based on a sample of 15,000 simulated patients. The robustness of these results is illustrated diagramatically in Section 6.

Section 3.12 of the draft ACD concludes that PSA results are not sufficiently reliable for decision making. The time pressures of reimplementing the model after the first ACM and in time for the ACD response, however, has meant we were unable to include the improved PSA results in this document. We are determined to provide the committee with PSA results sufficient for robust decision making prior to the second ACM meeting on the 8th June. We are grateful to the NICE project manager and ERG for their understanding of this situation.

3.6 Outstanding ERG preferences

<u>Annual BMI updates</u>

Section 1.5 of the ERG report noted the lack of an annual updating event for BMI as a validity issue for the submitted model, and this was mentioned in the NICE Lead Team Economic Presentation at the ACM on 6th April. The ERG were correct; the omission of annual BMI and time-to-event updates from the submitted model is explained by DICE run time, and a salient example of the need to reimplement the model in a faster platform.

Inclusion of annual model updates is important for good practice, and will allow the committee to confidently recommend NB32 for NHS patients on the basis of a robust and transparent economic analysis. Results from the Revised Base Case in Section 4 show that the introduction of annual updates reduces the estimated incremental QALY benefit of NB32. The implications of the changes to economic results should be considered; as stressed in the letter at the head of this response and in Sections 4 and 5, the results must be considered in the context of the systematic underlying

biases and inherent underestimation of the health and health care cost benefits of weight loss in the economic analysis against NB32.

Standard management healthcare resource assumptions

ERG amendment 7 removed the GP visit cost associated with annual assessment of people receiving standard management. We included this based on clinical advice on likely NHS practice,[8] but the ERG could not understand why this cost was applied. The assumption is not consequential and we therefore include amendment 7 in the Revised Base Case.

Approach to estimate 12-week orlistat treatment effectiveness

ERG amendment 3 attempted to implement use of a relative risk (RR) instead of mean difference to extrapolate the difference between treatments in change from baseline weight from the secondary to the primary assessment. The ERG introduced logic to change cells “I53:56” in woksheet “Efficacy” when control “ERG_weightchange_RR” is set to 1. However, cells “I53:56” do not feed through to model calculations, and so results were not affected.

We implemented the change the ERG had intended, but this introduced nonsensical model results for some patients, which we feel would surely have led the ERG to revise their preference. To illustrate, consider the primary assessment weight loss estimates produced for one simulated diabetic patient treated with orlistat when the RR approach is used:

- Weight loss at 16 weeks (NB32) – Responders: 5.3026%
- Weight loss at 56 weeks (NB32) - Primary assessment responders: 0.0237%
- Weight loss at 52 weeks (ORL; diabetic): 1.2223%

Using the ERG’s calculations, this orlistat patient would be sampled to achieve a weight loss at primary assessment of: 5.3026%*(1.2223%/0.0237%) = 273.0114%. Clearly this is impossible, and the intended amendment 3 of Table 3 was therefore not included in the Revised Base Case.

3.7 Correction of minor modelling errors

Some further minor technical fixes were applied within the model to ensure the ERG's preferred settings functioned as anticipated, as well as noting some modelling errors that were identified in the thorough reimplementation process. For example, the ERG's implementation of their preferred ITC analysis set did not apply to the specific cell ranges that informed the DICE model outputs. In addition, the application of the natural history model to predict BMI was applied inconsistently across different events; the equation includes a negative covariate for sex which should apply for male patients. For full clarity, a full description of the these minor technical corrections is provided to NICE as part of a technical reimplementation description, alongside this response and the reimplemented model containing Revised Base Case results.

3.8 The Revised Base Case

Table 4 summarises how the Revised Base Case differs from the CS model, referring the reader to the Section of the draft ACD each change corresponds to, and indicating whether the deviation from the CS base case is also a feature of the ERG base case.

Table 4: Summary of Differences between The Revised Base Case and The Company Submission Base Case

Feature of Revised Base Case that deviates from Company Submission Base Case assumptions	Present in ERG base case?	Relevant ACD Section
Subsequent bariatric surgery incorporated*	No	3.9
Reimplemented in efficient platform	No	3.10
Addresses concern over first-order uncertainty bias	No	3.11
Addresses concern over second-order uncertainty bias**	No	3.12
BMI natural history model calibrated to reflect trial baseline BMI	Yes	3.13
Baseline patient characteristics calibrated to reflect trial patients	Yes	3.13
ITT population from COR-I and COR-DM trials to inform treatment effectiveness analysis	Yes	3.14
Primary and secondary treatment phase TTD data from COR-I and COR-DM	No	3.14
Linear scaling assumption for orlistat TTD removed	Yes	3.14
Annual updating of BMI and TTE	No	Not mentioned
Standard management annual GP visit removed	Yes	Not mentioned
Correction of minor modelling errors	Yes	N/A

*Subsequent pharmacological treatment tested in a scenario

**PSA results with extended scope and sufficient PSA simulations to be shared with NICE in advance of 2nd ACM on 8th June 2017

Key: ACD, Appraisal Committee Document; ACM, Appraisal Committee Meeting; BMI, Body mass index; DM, diabetes mellitus; ERG, Evidence Review Group; GP, general practitioner; ITT, intention-to-treat; NICE, National Institute for Health and Care Excellence; PSA, Probabilistic sensitivity analysis; TTD, time-to-discontinuation; TTE, time-to-event.

4 Revised model results

Table 5 shows deterministic Revised Base Case results. Incremental results are relevant for those NHS patients whom could currently receive orlistat adjunct or SM alone. The pairwise comparison between NB32 and SM is salient for NHS patients whose treatment options are restricted to SM.

These deterministic results are based on a sample of 15,000 simulated patients. Figure 1, shown in the Appendix in Section 6 of this response, illustrates how mean incremental results change as the number of patient runs informing the analysis increases. Mean deterministic results based on 15,000 patient simulations are visibly robust to further increases in patient simulations.

The estimated lifetime health benefit of NB32 versus SM is 0.0433 QALYs or 0.0140 life years. We ask the committee to consider carefully the limitations of the analysis when interpreting these results. In particular:

- Only three obesity diseases are captured; MI, stroke and T2DM; when weight is a known risk factor for over 60 further health events[1], including numerous cancers[2, 3]
- Risk of death is assumed not to increase upon incidence of MI, stroke or T2DM
- The direct influence of BMI upon risk of death is incorporated for the first 15 years of the time horizon, but beyond this there is no assumed relationship between BMI and mortality risk

If data on *any* of these known limitations were incorporated into the analysis, the expected health benefit of NB32 could be better demonstrated. An additional 0.009 incremental QALY benefit would reduce the Revised Base Case ICER below £20,000. If each of the three limitations listed could be addressed to a satisfying level, the estimated health benefit of NB32 would be far greater than we can capture here. Taking the ignorance of 60 obesity diseases in isolation, if the relationship between weight loss and obesity risks could be fairly quantified, the NHS healthcare savings from disease prevention would very likely outweigh the £1,029 incremental cost of NB32 adjunct versus SM alone, and NB32 would be predicted to be a cost-saving, health improving resource.

Notably, Public Health England's "Weight management economic assessment tool", from which the analyses informing utility assumptions in this appraisal was sourced, considers the implications of BMI for both colorectal cancer and breast cancer incidence, and incorporates estimates of excess mortality from these diseases as well as heart attack, stroke and diabetes into assessments of the value of public health strategies to reduce obesity.[27]

Table 5: Revised base-case model results

Technologies	Total			Incremental			ICER	
	Costs	LYs	QALYs	Costs	LYs	QALYs	Versus baseline [SM]	Incremental
Incremental								
SM	£6,502	33.9109	13.6300					
ORL	£6,802	33.9225	13.6698	£300	0.0116	0.0398	£7,536	£7,536
NB32	£7,531	33.9243	13.6734	£729	0.0018	0.0035	£23,750	£207,274
Pairwise								
SM	£6,502	33.9109	13.6300					
NB32	£7,531	33.9243	13.6734	£1,029	0.0134	0.0433		£23,750
ORL	£6,802	33.9225	13.6698					
NB32	£7,531	33.9243	13.6734	£729	0.0018	0.0035		£207,274
Key: ICER, incremental cost-effectiveness ratio; LYs, life years; NB32, naltrexone bupropion; ORL, orlistat; QALYs, quality-adjusted life years; TTD, time to treatment discontinuation.								

Table 6 and Table 7 show results for baseline non-T2DM and baseline T2DM patient subgroups, respectively. Reflecting previous iterations of results for these subgroups, NB32 is estimated to be more effective and cost-effective in the group without T2DM at treatment initiation, and may be a particularly valuable treatment option for eligible patients without T2DM for whom orlistat is not a treatment option.

Table 6: Revised base-case model results, non-T2DM subgroup

Technologies	Total			Incremental			ICER	
	Costs	LYs	QALYs	Costs	LYs	QALYs	Versus baseline [SM]	Incremental
Incremental								
SM	£4,300	34.0375	14.0335					
ORL	£4,572	34.0487	14.0669	£272	0.0111	0.0334	£8,153	£8,153
NB32	£5,311	34.0528	14.0797	£738	0.0041	0.0128	£21,897	£57,899
Pairwise								
SM	£4,300	34.0375	14.0335					
NB32	£5,311	34.0528	14.0797	£1,011	0.0152	0.0462		£21,897
ORL	£4,572	34.0487	14.0669					
NB32	£5,311	34.0528	14.0797	£738	0.0041	0.0128		£57,899
Key: ICER, incremental cost-effectiveness ratio; LYs, life years; NB32, naltrexone bupropion; ORL, orlistat; QALYs, quality-adjusted life years; TTD, time to treatment discontinuation; T2DM, Type 2 Diabetes Mellitus.								

Table 7: Revised base-case model results, T2DM subgroup

Technologies	Total			Incremental			ICER	
	Costs	LYs	QALYs	Costs	LYs	QALYs	Versus baseline [SM]	Incremental
Incremental								
SM	£11,435	33.5577	12.7100					
NB32	£12,467	33.5659	12.7496	£1,032	0.0081	0.0396	£26,049	£26,049
ORL	£11,785	33.5688	12.7639	-£681	0.0030	0.0143	£6,507	Dominant
Pairwise								
SM	£11,435	33.5577	12.7100					
NB32	£12,467	33.5659	12.7496	£1,032	0.0081	0.0396		£26,049
ORL	£12,467	33.5659	12.7496					
NB32	£11,785	33.5688	12.7639	-£681	0.0030	0.0143		Dominant
Key: ICER, incremental cost-effectiveness ratio; LYs, life years; NB32, naltrexone bupropion; ORL, orlistat; QALYs, quality-adjusted life years; TTD, time to treatment discontinuation.								

Table 8 shows how results change with each deviation away from the ERG base case and towards the Revised Base Case. While there is very little change in estimated incremental costs, the estimated incremental QALYs associated with NB32 versus SM have fallen substantially, from 0.100 QALYs to 0.043 QALYs, causing the estimated ICER for this comparison to increase from just over £10,500 to just under £23,750. The comparison to orlistat is also less favourable than in the ERG base case, with the extremently senistive ICER versus orlistat increasing from less than £50,000 (first-order uncertainty around ERG base case results considered), to over £200,000. The limitations of this comparison, stressed in Section 1 of this document, should be considered carefully when interpreting results.

The incorporation of an annual BMI and TTE updating event to the analysis had by far the most profound impact on the results. Incorporating annual updates led to a reduction in total life years across treatment arms, and a reduction in estimated incremental life years for both orlistat and NB32 versus SM. This contributes to a reduction in total QALYs across arms, and a reduction in estimated incremental QALYs for both NB32 and orlistat versus SM, and for NB32 versus orlistat. The change in total and incremental health outcome results unfold from death occurring earlier and utility estimates being updated for natural increases in age and BMI more regularly; the incremental gains are diminished as totals are diminished, and for some patients death now precedes incidence of a health event, whereas without updating this health event was predicted to occur before death, with consequences for incremental QALY estimates.

Aside from the incorporation of an annual updating event, implementing the minor technical corrections described in Section 3.7 of this response had a notable effect upon results from the comparison to orlistat. Ensuring the ERG's ITC preference were implemented as they intended reduced the estimated clinical benefit of NB32 versus orlistat, causing the extremely sensitive ICER estimate in this comparisoon to rise from less than £42,000 to more then £90,000. Other changes from the ERG base case had more marginal influence upon results.

Table 8: Overview of Route from the ERG Base Case to the Revised Base Case

Technologies	Total			Incremental			ICER
	Costs	LYs	QALYs	Costs	LYs	QALYs	
(1) = "ID757 Naltrexone - bupropion HE ERG_base-case 090317 LG [ACIC]", n=1,000							
SM	£5,964	34.4406	15.1111				
NB32	£7,017	34.4814	15.2113	£1,053	0.0408	0.1002	£10,510
ORL	£6,275	34.4760	15.1950				
NB32	£7,017	34.4814	15.2113	£742	0.0054	0.0162	£45,694
(2) = Reimplementation of (1) in VBA, n=1,000							
SM	£5,864	34.4920	15.1339				
NB32	£6,908	34.5198	15.2324	£1,044	0.0278	0.0985	£10,600
ORL	£6,166	34.5145	15.2160				
NB32	£6,908	34.5198	15.2324	£742	0.0053	0.0163	£45,488
(3) = (2), n=15,000							
SM	£6,016	34.8242	15.2416				
NB32	£7,052	34.8589	15.3394	£1,036	0.0347	0.0978	£10,594
ORL	£6,305	34.8525	15.3214				
NB32	£7,052	34.8589	15.3394	£746	0.0064	0.0180	£41,552
(4) = (3) with minor technical corrections (see Section 3.7), n=15,000							
SM	£6,106	34.4682	14.9644				
NB32	£7,148	34.4974	15.0555	£1,042	0.0293	0.0911	£11,435
ORL	£6,404	34.4934	15.0473				
NB32	£7,148	34.4974	15.0555	£744	0.0041	0.0082	£90,640
(5) = (4) with ERG amendments 1, 4 and 7 removed (see Sections 3.3 and 3.6), n=15,000							
SM	£6,190	34.3812	14.8765				
NB32	£7,225	34.4255	14.9807	£1,036	0.0443	0.1043	£9,935
ORL	£6,483	34.4196	14.9707				
NB32	£7,225	34.4255	14.9807	£742	0.0059	0.0100	£74,188
(6) = (5) with annual updating event implemented (see Section 3.6), n=15,000							
SM	£6,507	33.9113	13.6316				
NB32	£7,562	33.9250	13.6765	£1,056	0.0137	0.0449	£23,532
ORL	£6,815	33.9230	13.6724				
NB32	£7,562	33.9250	13.6765	£747	0.0019	0.0041	£184,318
(7) = (6) with subsequent bariatric surgery incorporated (see Section 3.4), n=15,000							
SM	£6,512	33.9113	13.6316				
NB32	£7,568	33.9250	13.6764	£1,056	0.0137	0.0449	£23,533
ORL	£6,820	33.9230	13.6724				
NB32	£7,568	33.9250	13.6764	£747	0.0019	0.0041	£184,309
(8) = (7) with TTD from COR-I and COR-DM only (see Section 3.2), n=15,000 = Revised Base Case							
SM	£6,502	33.9109	13.6300				
NB32	£7,531	33.9243	13.6734	£1,029	0.0134	0.0433	£23,750
ORL	£6,802	33.9225	13.6698				
NB32	£7,531	33.9243	13.6734	£729	0.0018	0.0035	£207,274

Key: DM, diabetes mellitus; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; LY, life year; NB, naltrexone-bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management; TTD, time to discontinuation; VBA, Visual Basic for Applications®

Table 9 shows results from a scenario where NB32 is compared to SM in patients who have failed to achieve 5% weight loss from 12 weeks of adjunctive orlistat treatment. As expected given the necessary assumptions of the analysis, the incremental results from this scenario are very similar to the Revised Base Case results in Table 5. Given the implicit bias against NB32 in these analyses, the

innovative nature of NB32 treatment in this indication, and the unmet need faced by these patients, NB32 is a valuable treatment for patients who have attempted to lose weight with orlistat but failed to meet 12-week weight loss requirements for treatment continuation.

Table 9: Scenario analysis, exploring NB32 versus SM for those who fail orlistat

Technologies	Total			Incremental			ICER
	Costs	LYs	QALYs	Costs	LYs	QALYs	
SM	£6,527	33.9238	13.6404				
NB32	£7,557	33.9382	13.6845	£1,030	0.0144	0.0442	£23,324
Key: ICER, incremental cost-effectiveness ratio; LYs, life years; NB32, naltrexone bupropion; ORL, orlistat; QALYs, quality-adjusted life years.							

We look forward to providing results from PSA of the Revised Base Case as soon as possible, and to the 2nd ACM on 8th June. We hope that the effort we have made to both reimplement the model in a suitably efficient platform and provide the committee with transparent and open analyses designed to meet committee preferences will allow the committee to recommend NB32 as an innovative treatment option alongside SM, for those patients who have the will to reduce their weight to healthy levels, but require additional pharmacological support to achieve their goal.

5 References

1. Yumuk V, Tsigos C, Fried M, et al. European Guidelines for Obesity Management in Adults. *Obes Facts*. 2015; 8(6):402-24.

2. Dasika G and Nielsen SK. The increasing cost burden of obesity related cancer to NHS in the UK. ISPOR 18th Annual European Congress. Milan, Italy. November 2015 2015. PCN110.

3. World Obesity Federation. Estimates of relative risk of disease per unit of BMI above 22kg/m2. 2015. Available at: http://www.worldobesity.org/what-we-do/policy-prevention/projects/eu-projects/dynamohiaproject/estimatesrrperunitbmi/. Accessed: 23 November 2016.

4. Sullivan PW, Ghushchyan VH and Ben-Joseph R. The impact of obesity on diabetes, hyperlipidemia and hypertension in the United States. *Qual Life Res*. 2008; 17(8):1063-71.

5. Cuesta-Vargas AI and González-Sánchez M. Obesity effect on a multimodal physiotherapy program for low back pain suffers: Patient reported outcome. *J Occup Med Toxicol*. 2013; 8(1).

6. Foster SA, Hambright DS, Antoci V, et al. Effects of obesity on health related quality of life following total hip arthroplasty. *J Arthroplasty*. 2015; 30(9):1551-4.

7. Knutsson B, Michaëlsson K and Sandén B. Obesity is associated with inferior results after surgery for lumbar spinal stenosis: A study of 2633 patients from the Swedish spine register. *Spine*. 2013; 38(5):435-41.

8. Orexigen. Clinical validation meeting for Mysimba® NICE submission. 29 September 2016. Data on file.

9. National Institute of Health and Care Excellence (NICE). CG189: Obesity: identification, assessment and management. 2014. Available at: https://www.nice.org.uk/guidance/cg189/chapter/1-recommendations. Accessed: 12 October 2016.

10. Derosa G, Maffioli P, Salvadeo SA, et al. Comparison of orlistat treatment and placebo in obese type 2 diabetic patients. *Expert Opin Pharmacother*. 2010; 11(Issue):1971-82.

11. Karhunen L, Franssila-Kallunki A, Rissanen P, et al. Effect of orlistat treatment on body composition and resting energy expenditure during a two-year weight-reduction programme in obese Finns. *IntJ Obes Relat Metab Disord*. 2000; 24(12):1567-72.

12. Mathus-Vliegen EM, Leeuwen MLI-v and Bennink RJ. Influences of fat restriction and lipase inhibition on gastric emptying in obesity. *Int J Obes (Lond)*. 2006; 30(Issue):1203-10.

13. Reaven G, Segal K, Hauptman J, et al. Effect of orlistat-assisted weight loss in decreasing coronary heart disease risk in patients with syndrome X. *Am J Cardiol*. 2001; 87(7):827-31.

14. Astrup A, Carraro R, Finer N, et al. Safety, tolerability and sustained weight loss over 2 years with the once-daily human GLP-1 analog, liraglutide. *Int J Obes (Lond)*. 2012; 36(Issue):843-54.

15. Berne C. A randomized study of orlistat in combination with a weight management programme in obese patients with Type 2 diabetes treated with metformin. *Diabet Med*. 2005; 22(5):612-8.

16. Kelley DE, Bray GA, Pi-Sunyer FX, et al. Clinical efficacy of orlistat therapy in overweight and obese patients with insulin-treated type 2 diabetes: A 1-year randomized controlled trial. *Diabetes Care*. 2002; 25(6):1033-41.

17. Miles JM, Leiter L, Hollander P, et al. Effect of orlistat in overweight and obese patients with type 2 diabetes treated with metformin.[Erratum appears in Diabetes Care. 2002 Sep;25(9):1671.]. *Diabetes Care*. 2002; 25(Issue):1123-8.

18. Caro JJ. Discretely Integrated Condition Event (DICE) Simulation for Pharmacoeconomics. *PharmacoEconomics*. 2016; 34(7):665-72.

19. Food and Drug Administration. Guidance for Industry Developing Products for Weight Management: Draft Guidance. Revision 1 ed. Rockville, MD2007.

20. Ara R, Blake L, Gray L, et al. What is the clinical effectiveness and cost-effectiveness of using drugs in treating obese patients in primary care? A systematic review. *Health Technol Assess*. 2012; 16(5):iii-xiv, 1-195.

21. British Medical Journal (BMJ). Press Release: NHS needs to perform more weight loss surgery to curb the obesity epidemic, argue experts 2016. Available at: http://www.bmj.com/company/wp-content/uploads/2016/05/NHS-weight-loss-surgery.pdf. Accessed: 22 May 2017.

22. Picot J, Jones J, Colquitt JL, et al. The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation. *Health Technol Assess*. 2009; 13(41):1-190, 215-357, iii-iv.

23. NHS Choices. Weight loss surgery. 2017. Available at: http://www.nhs.uk/Conditions/weight-loss-surgery/Pages/Introduction.aspx. Accessed: 22 May 2017.

24. Madura JA and DiBaise JK. Quick fix or long-term cure? Pros and cons of bariatric surgery. *F1000 Medicine Reports*. 2012; 4:19.

25. Kellow J. Weight Loss Surgery Questions + Answers. Available at: http://www.weightlossresources.co.uk/weight_loss/surgery.htm. Accessed: May 2017.

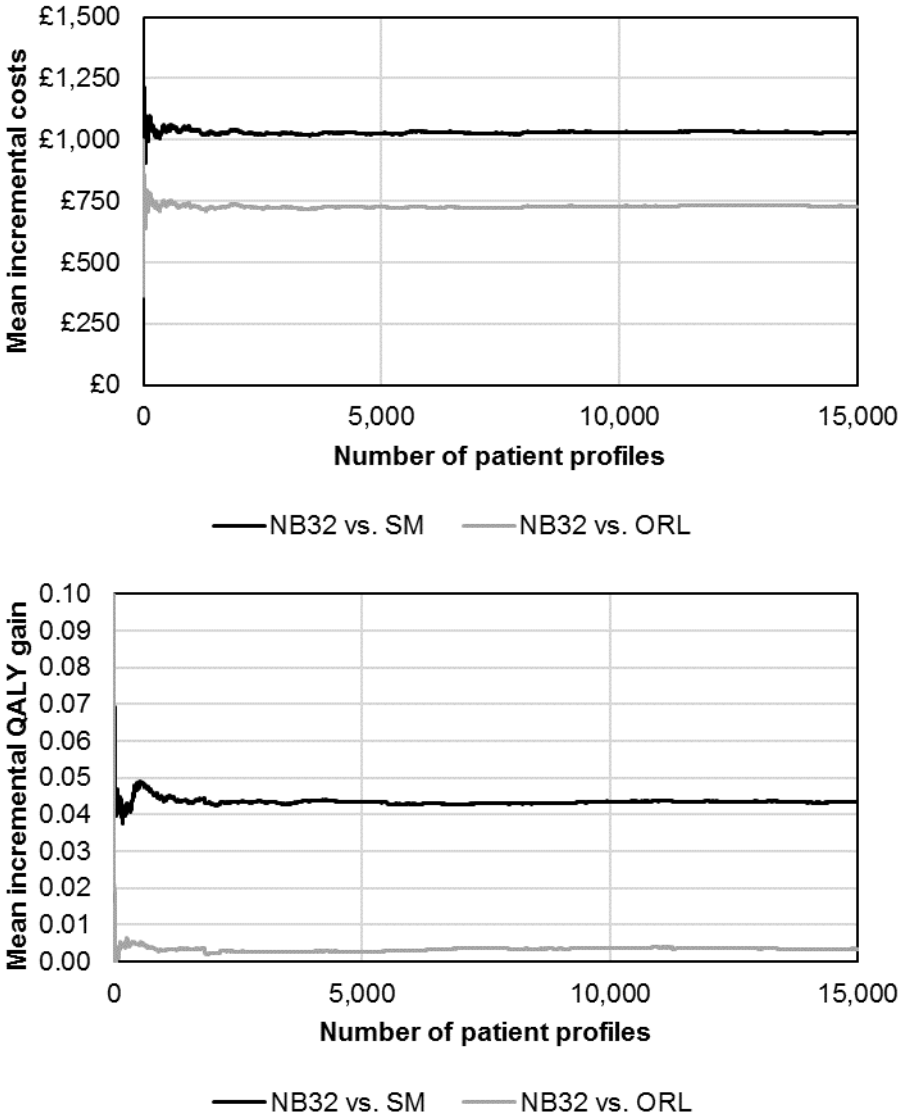
26. Department of Health. NHS Reference Costs. 2015. Available at: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015. Accessed: 22 Sept 2016.

27. Copley V. *User Guide: Weight Management Economic Assessment Tool Version 2*. Oxford: Public Health England, Obesity Risk Factors Intelligence, 2016.

6 Appendix

Figure 1 show how mean estimated incremental costs and QALYs change as the number of patient runs informing the analysis increases, for pairwise comparisons to SM and orlistat, respectively. Mean deterministic results based on 15,000 patient simulations are visibly robust to further increases in patient simulations.

Figure 1: Stabilisation of mean incremental costs and QALYs by number of patient simulations



Key: NB32, naltrexone-bupropion; QALY, quality-adjusted life year; SM, standard management.

Explanation of re-implementation and validation of CE model for NB32 for obesity (ID757)

Overview of transfer of calculations from "ID757 Naltrexone - bupropion HE ERG_base-case 090317 LG [ACIC]" ("DICE model") to VBA logic

To transfer the model into VBA code, the following process was undertaken:

- The overarching DICE technique (that is, the layout of constants, profiles, events etc.) was moved into VBA form.
- Constants and profiles were read into VBA as arrays (equivalent to tables in Excel, but the data in arrays are held in memory and therefore calculations using arrays may be carried out more quickly). At the beginning of the run, patient profiles are read in via the array and default values are assigned (such as setting the time to stroke as 999999 until calculated).
- A *"DICE_run"* macro was written that is equivalent to the derivation of the next event time within the DICE model. In this macro, each event is compared to the next event time until a match is obtained. Following a match, a separate macro is triggered for the event under consideration.
- Each event has two macros associated with it. First, a *"DICE_events_..."* macro detailing the events and conditions that are updated in the same layout as per the original DICE model. Secondly, any custom functions required are incorporated into a *"DICE_Func_..."* macro (for example, the derivation of BMI using the natural history model requires multiple lines of calculation, and is therefore presented as a custom function).
- Outputs from the VBA model are in the same format as per the original DICE model, though are now pasted onto the *"DICE_Output"* sheet.

Throughout the transfer of equations into VBA, some functions were more complex to implement than others, due to the differences in logic between VBA code and Excel worksheet functions. Some mathematical functions such as additions and multiplications were simple to re-implement via a copy and paste. Other functions such as *"IF"* statements were slightly more cumbersome to re-implement – nested *"IF"* statements in VBA code require separation over several lines of code.

The next level of complexity considers those functions that do not have a VBA counterpart – for example, a vertical lookup (or *"VLOOKUP"*) function. This type of function requires VBA code to state that it is a worksheet function by including the text *"Application.WorksheetFunction."* ahead of the name of the function. Finally, some DICE-specific features required complete re-structuring, as the equivalent function cannot be considered in VBA. For example, in the DICE model the next event was selected using *"live"* values and a *"live" "MIN"* function. In VBA, the selection of an event had to be coded as a loop, where each event was assessed as a candidate for the next event with the smallest of these times selected based on a conditional subroutine called to actively calculate the minimum of the derived values.

Issues identified during re-implementation of "ID757 Naltrexone - bupropion HE ERG_base-case 090317 LG [ACIC]"

In addition, when transferring model equations into VBA some minor errors were identified in the original model. The identification of these errors was discussed briefly in Section 3.7 of the ACD response and were:

- The ERG's changes to the DICE model included their preferred ITC estimates, but these estimates did not apply to the correct cells (cells I57 and I58 on the sheet *"Efficacy"* in the DICE model did not link to further calculations).
- The application of the natural history model for BMI includes a BMI decrement for male patients. Certain model equations in the DICE model erroneously considered the decrement to apply for females instead (cells D419 and I416 on the sheet *"DICE equations"* in the ERG-amended DICE model incorporated these errors).
- There was an error in the application of the Tobit utility equation for the calculation of utility upon a patient entering the model. This error was also present in the manufacturer-submitted DICE model (cell D148 on the sheet *"DICE equations"* applied this model as per a standard linear regression, which is incorrect).

Changes to fix identified issues

A switch has been incorporated into the re-implemented model called *"ERG_include_error"* on the sheet *"ERG"*. This switch enables and disables the errors described above (1=enabled). A further switch *"ERG_fix"* ensures outcomes for diabetic and non-diabetic patients are as intended – the ERG model accounts for diabetic status at baseline within its calculations, however for efficiency the VBA model considers only the *"diabetic_status"* condition. Therefore, the *"ERG_fix"* switch should always be set to 1.

Further switches in the model consider the scenarios presented in response to the ACD, where a value of 1 is equivalent to turning on the switch. It should be noted that the switch for *"ERG_weightchange_RR"* should permanently be set to 0, as the implementation of this switch introduces model errors.

Validation of re-implementation of the model in VBA code

To validate the outcomes of the VBA model compared with the ERG model, a number of steps were taken. Firstly, a thorough comparison of the equations written into the VBA code versus those presented in the ERG model was undertaken. This process was facilitated by copying and pasting each equation into VBA, converting the equation into VBA code (e.g. implementing *"Application.WorksheetFunction."* where required) and sense checking where possible that outputs were aligned as expected (e.g. calculating the Tobit model utility at baseline for patients using the VBA code function and comparing the outcome to the value obtained in the Excel calculations).

Following careful implementation, a model run of 1,000 patients was conducted (as per the ERG base case). The cost-effectiveness results from both models were compared, and were deemed suitably comparable to then compare individual patient profile results and obtain *"match rates"* based on how similar the results across both models were.

In the majority of cases, results matched near-perfectly with some differences noted mainly in the estimation of standard management and condition management costs. These differences were investigated and resolved in an iterative process until no further improvements could be made, and the differences between results were sufficiently small for them to be considered inconsequential for inference.

Brief instructions for re-running results in the VBA model

To run the base-case deterministic analysis in the VBA model, the following steps should be taken:

- Click on the *"Run quicker VBA model"* button on the *"Base case results"* sheet
- Choose the desired number of patient profiles (base case = 15,000)

Deterministic model execution less than 10 minutes to run for a sample of 15,000 simulated patients, on a laptop with 8GB RAM and an i7 core processer. To run1,000 PSA iterations would therefore take around 7 days of continuous processing on a single machine. Therefore, it is advised that if repeating a PSA run, the PSA can be staggered over a number of machines (running 250 across 4 machines, 1000 PSA iterations should take less than 2 days to complete). To run the PSA, the following steps should be taken:

- Click on the *"Run quicker PSA"* button on the *"DICE_Output_PSA"* sheet
- Choose the desired number of patient profiles (deterministic base case = 15,000)
- Choose the starting number to run from using this machine (if the PSA is to be run on one machine, enter 1, else the user is able to stagger results across a number of machines if run time is a limitation).
- Choose the number of PSA iterations

Orexigen provision of updated probabilistic sensitivity analysis results to:

National Institute for Health and Care Excellence

Naltrexone-bupropion (prolonged release) for managing overweight and obesity [ID757]

June 2017

Dear Committee Members and the Evidence Review Group (ERG),

As stated in our response to the Appraisal Consultation Document (ACD), we are grateful to the NICE project manager and ERG for their understanding of the time pressures of reimplementing the model after the first Appraisal Consultation Meeting (ACM) and in time for the ACD response. The results of the probabilistic sensitivity analysis (PSA) have now been produced and are discussed in this document.

Yours sincerely,

Hans-Joerg Fugel

VP Market Access Europe,

Orexigen Therapeutics

1. Approach taken to produce probabilistic sensitivity analysis results

At the appraisal committee meeting on 06 April 2017 and in the ACD a number of issues were raised with the original PSA. These issues were summarised in Section 3.12 of the ACD:

"The ERG explained that the company did not run enough iterations to produce stable results for the probabilistic sensitivity analysis (PSA). The committee was aware that past PSAs run over 1,000 iterations to produce results, but the company ran only 500. The committee also noted that the company had not included important parameters in the PSA, which are subject to great uncertainty (time to treatment discontinuation [TTD], natural history of BMI model, and obesity-related events). The committee concluded that the PSA results were not sufficiently reliable for decision-making."

The ACD also commented that the number of patient simulations used to produce deterministic (and by extension probabilistic) results was considered *"too few"* (Section 3.11). To address these concerns, the PSA was re-implemented using Visual Basic for Applications (VBA) logic, as communicated in our 30 May 2017 ACD Response. This alleviated the computational burden of producing PSA results, and therefore made the production of PSA results with a large enough number of patient simulations and PSA iterations possible.

The revised PSA considers 1,000 PSA iterations, each of which was ran using 15,000 patient simulations as per the revised deterministic base-case analysis presented in response to the ACD. The mean probabilistic result was shown to be stable for the comparison of NB32 and standard management, as shown in Figure 1. For the comparison of NB32 and orlistat, the ICER was less stable but showed a reasonable degree of convergence by 1,000 iterations as shown in Figure 2.

To address the ACD concern that important parameters were omitted from the PSA, the model was first updated to incorporate the standard deviation (SD) of the estimates used to inform the body mass index (BMI) trajectory natural history model. Though variance-covariance matrices were not reported by Ara *et al*, the 95% confidence intervals (CI's) reported for each parameter were used to meet the committee's and ERG's request to incorporate uncertainty around these parameters. Each BMI trajectory model parameter is assumed to be independently and normally distributed in the PSA.

The other parameters flagged within the ACD that were not explored in the original PSA were those relating to the time to treatment discontinuation (TTD) and the time to obesity-related events. The TTD estimates within the model were derived from the Kaplan-Meier

(KM) estimates from the COR trial programme. As part of the response to the ACD, only the COR-I and COR-DM studies were deemed appropriate for use within the model. To incorporate the uncertainty of the KM curves, the area under the curve was obtained using the mean, lower and upper estimates of the KM. From these values, an estimate of the standard error (SE) of the curve was obtained by using the Solver® function in Excel® to back-calculate the estimated SE based on the Greenwood estimate of the 95% CI around the KM.[1]

For the time to all-cause mortality (ACM), myocardial infarction (MI), stroke and onset type-2 diabetes mellitus (T2DM); uncertainty was incorporated within PSA by using a "scaling factor". In the deterministic analysis, this factor was set to 1, whereas in PSA a random number was drawn from a normal distribution with mean 1 and SD 0.1 (that is, 10% of the mean value). The factor was used to vary the predicted times to events while maintaining the correlation between specific parameters (e.g. the ln(λ) contribution of BMI^2). As stated in the original company submission, incorporation of the estimated error structure time-to-event models from Ara *et al* is not possible, due to the absence of the relevant variance-covariance matrices from their *Health Technology Assessment* publication. However, the use of the "scaling factors" allows for some variation to be explored within the PSA.

2. Summary of parameters explored in probabilistic sensitivity analysis

Table 1: Updated summary of variables applied in the economic model

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
Discount rate for costs	3.5%	Varied between 0% and 6% in OWSA; Varied as 3.5%, 1.5% and 0% in scenario analysis; otherwise fixed.	Section 5.2
Discount rate for QALYs	3.5%		
Discount rate for LYs	0.0%		
Sample age (mean in base case)	47.00	Patient profile specific - sampled at baseline and therefore not explored within sensitivity analysis.	Section 5.3.1
Sample height (female; mean in base case)	1.64		
Sample height (male; mean in base case)	1.78		
Proportion of patients female	79.0%		
Proportion of patients male	21.0%		
Proportion of patients with T2DM	33.2%		
Proportion of patients without T2DM	66.8%		
Proportion of patients who currently smoke	7.0%		
Proportion of patients who previously smoked	54.0%		
Proportion of patients who have never smoked	39.0%		
Proportion of patients who receive insulin (if diabetic)	33.3%		
Proportion of patients who receive anti-hypertensive medication	17.0%		
Proportion of patients who receive statins	79.3%		
Proportion of patients with history of angina	0.0%		
Proportion of patients with history of non-Type 2 diabetes mellitus	0.0%		
Proportion of patients who receive aspirin	10.7%		
Drug cost of NB32 8mg/90mg, 112 tab pack	£73.00	Fixed	Section 5.5.2
Cost of ORL per pack (84 capsules) - generic	£18.44		
Administration cost NB32	£0.00	Fixed at zero	
Administration cost ORL	£0.00		
Administration cost SM	£0.00		
GP visit	£44.00	Gamma, SE assumed 10% of the mean	Section 5.5.2
Nurse visit	£14.47		
Blood test	£3.01		
MI (Year 1)	£4,210.75		Section 5.5.3

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
MI (Year 1+)	£345.91		
Stroke (Year 1)	£9,482.78		
Stroke (Year 1+)	£2,664.16		
T2DM (Year 1)	£347.57	Triangular, bounds from source	
T2DM (Year 1+)	£347.57		
Fatal stroke	£8,671.94	Gamma, SE assumed 10% of the mean	
Fatal MI	£1,390.80		
Probability CVD-related mortality	31.0%	Beta, assumed n=1000, SD 10% of the mean	
Probability MI mortality	43.1%	Dirichlet, assumed n=1,000	
Probability stroke mortality	32.9%		
Probability other CVD mortality	24.0%		
Outpatient: 300: General Medicine	£158.43	Gamma, SE assumed 10% of the mean	Section 5.5.4
Outpatient: 502: Gynaecology	£132.75		
Outpatient: 301: Gastroenterology	£135.18		
Outpatient: 710: Adult Mental Illness	£241.52		
Outpatient: 120: ENT	£94.36		
Outpatient: 400: Neurology	£175.76		
Tobit model BMI coefficient	0.059	Multi-variate normal distribution, variance-covariance matrix from source	Section 5.4.4
Tobit model BMI2 coefficient	-0.002		
Tobit model BMI3 coefficient	0.000		
Tobit model Age coefficient	-0.004		
Tobit model Female coefficient	-0.041		
Tobit model Stroke coefficient	-0.183		
Tobit model MI coefficient	-0.161		
Tobit model Cancer coefficient	-0.164		
Tobit model T2DM coefficient	-0.111		
Tobit model Constant coefficient	0.673		
OLS model BMI coefficient	0.033	Multi-variate normal distribution, variance-covariance matrix from source	
OLS model BMI2 coefficient	-0.001		
OLS model BMI3 coefficient	0.000		
OLS model Age coefficient	-0.002		
OLS model Female coefficient	-0.023		
OLS model Stroke coefficient	-0.127		
OLS model MI coefficient	-0.119		

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
OLS model Cancer coefficient	-0.109		
OLS model T2DM coefficient	-0.078		
OLS model Constant coefficient	0.658		
Weight regain time (years)	3 years	Varied ±10% of the mean in OWSA, explored within scenario analysis, otherwise fixed.	Section 5.2.4.1
NB32 primary phase TTD adjustment factor	1	Scaling factor obtained using area under the curve. Varied using a normal distribution with mean 1 and SD specific to the curve. See Section 1 of this document for further details.	
NB32 secondary phase TTD adjustment factor	1		
NB32 tertiary phase TTD adjustment factor	1		
SM TTD adjustment factor	1		
ACM - Diabetic - ln(λ) - BMI	0.27700	Scaling factor obtained for the predicted times to events using a normal distribution with mean 1 and SD 0.1. See Section 1 of this document for further details.	Section 5.3.4.2
ACM - Diabetic - ln(λ) - BMI2	-0.00338		
ACM - Diabetic - ln(λ) - BMI3	0.00000		
ACM - Diabetic - ln(λ) - BMI4	0.00000		
ACM - Diabetic - ln(λ) - Age	0.09000		
ACM - Diabetic - ln(λ) - Age2	0.00000		
ACM - Diabetic - ln(λ) - Age3	0.00000		
ACM - Diabetic - ln(λ) - Age4	0.00000		
ACM - Diabetic - ln(λ) - Age5	0.00000		
ACM - Diabetic - ln(λ) - Sex	0.41000		
ACM - Diabetic - ln(λ) - Aspirin treatment	-0.02600		
ACM - Diabetic - ln(λ) - Insulin treatment	-0.10300		
ACM - Diabetic - ln(λ) - Statin treatment	-0.80600		
ACM - Diabetic - ln(λ) - BP drug treatment	-0.20800		
ACM - Diabetic - ln(λ) - Ex-smoker	-0.13000		
ACM - Diabetic - ln(λ) - Smoker	0.32600		
ACM - Diabetic - ln(λ) - Constant	-17.25800		
ACM - Diabetic - ln(γ) - BMI	0.00000		
ACM - Diabetic - ln(γ) - BMI2	0.00000		
ACM - Diabetic - ln(γ) - Age	0.00000		
ACM - Diabetic - ln(γ) - Age2	0.00000		
ACM - Diabetic - ln(γ) - Sex	0.00000		
ACM - Diabetic - ln(γ) - Aspirin treatment	0.00000		
ACM - Diabetic - ln(γ) - Statin treatment	0.00000		
ACM - Diabetic - ln(γ) - BP drug treatment	0.00000		

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
ACM - Diabetic - ln(γ) - Ex-smoker	0.00000		
ACM - Diabetic - ln(γ) - Smoker	0.00000		
ACM - Diabetic - ln(γ) - Constant	0.78500		
ACM - Non-diabetic - ln(λ) - BMI	6.12300		
ACM - Non-diabetic - ln(λ) - BMI2	-0.21400		
ACM - Non-diabetic - ln(λ) - BMI3	0.00343		
ACM - Non-diabetic - ln(λ) - BMI4	-0.00002		
ACM - Non-diabetic - ln(λ) - Age	0.07900		
ACM - Non-diabetic - ln(λ) - Age2	0.00040		
ACM - Non-diabetic - ln(λ) - Age3	0.00000		
ACM - Non-diabetic - ln(λ) - Age4	0.00000		
ACM - Non-diabetic - ln(λ) - Age5	0.00000		
ACM - Non-diabetic - ln(λ) - Sex	1.41100		
ACM - Non-diabetic - ln(λ) - Aspirin treatment	-0.62000		
ACM - Non-diabetic - ln(λ) - Insulin treatment	0.00000		
ACM - Non-diabetic - ln(λ) - Statin treatment	-1.30100		
ACM - Non-diabetic - ln(λ) - BP drug treatment	-0.79700		
ACM - Non-diabetic - ln(λ) - Ex-smoker	-1.19600		
ACM - Non-diabetic - ln(λ) - Smoker	0.33800		
ACM - Non-diabetic - ln(λ) - Constant	-80.78100		
ACM - Non-diabetic - ln(γ) - BMI	-0.08500		
ACM - Non-diabetic - ln(γ) - BMI2	0.00104		
ACM - Non-diabetic - ln(γ) - Age	-0.00200		
ACM - Non-diabetic - ln(γ) - Age2	0.00000		
ACM - Non-diabetic - ln(γ) - Sex	-0.10800		
ACM - Non-diabetic - ln(γ) - Aspirin treatment	0.09800		
ACM - Non-diabetic - ln(γ) - Statin treatment	0.09500		
ACM - Non-diabetic - ln(γ) - BP drug treatment	0.11400		
ACM - Non-diabetic - ln(γ) - Ex-smoker	0.15600		
ACM - Non-diabetic - ln(γ) - Smoker	0.04100		
ACM - Non-diabetic - ln(γ) - Constant	2.60000		
MI (non-fatal) - Diabetic - ln(λ) - BMI	-0.00300		
MI (non-fatal) - Diabetic - ln(λ) - BMI2	0.00000		
MI (non-fatal) - Diabetic - ln(λ) - BMI3	0.00000		

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
MI (non-fatal) - Diabetic - ln(λ) - BMI4	0.00000		
MI (non-fatal) - Diabetic - ln(λ) - Age	0.04400		
MI (non-fatal) - Diabetic - ln(λ) - Age2	0.00000		
MI (non-fatal) - Diabetic - ln(λ) - Age3	0.00000		
MI (non-fatal) - Diabetic - ln(λ) - Age4	0.00000		
MI (non-fatal) - Diabetic - ln(λ) - Age5	0.00000		
MI (non-fatal) - Diabetic - ln(λ) - Sex	1.20600		
MI (non-fatal) - Diabetic - ln(λ) - Aspirin treatment	0.63800		
MI (non-fatal) - Diabetic - ln(λ) - Insulin treatment	0.60900		
MI (non-fatal) - Diabetic - ln(λ) - Statin treatment	2.11900		
MI (non-fatal) - Diabetic - ln(λ) - BP drug treatment	0.98000		
MI (non-fatal) - Diabetic - ln(λ) - Ex-smoker	-0.47800		
MI (non-fatal) - Diabetic - ln(λ) - Smoker	0.32900		
MI (non-fatal) - Diabetic - ln(λ) - Constant	-11.92100		
MI (non-fatal) - Diabetic - ln(γ) - BMI	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - BMI2	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - Age	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - Age2	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - Sex	-0.20700		
MI (non-fatal) - Diabetic - ln(γ) - Aspirin treatment	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - Statin treatment	-0.50600		
MI (non-fatal) - Diabetic - ln(γ) - BP drug treatment	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - Ex-smoker	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - Smoker	0.00000		
MI (non-fatal) - Diabetic - ln(γ) - Constant	0.79500		
MI (non-fatal) - Non-diabetic - ln(λ) - BMI	0.03000		
MI (non-fatal) - Non-diabetic - ln(λ) - BMI2	0.00000		
MI (non-fatal) - Non-diabetic - ln(λ) - BMI3	0.00000		
MI (non-fatal) - Non-diabetic - ln(λ) - BMI4	0.00000		
MI (non-fatal) - Non-diabetic - ln(λ) - Age	-1.14100		
MI (non-fatal) - Non-diabetic - ln(λ) - Age2	0.06000		
MI (non-fatal) - Non-diabetic - ln(λ) - Age3	-0.00136		
MI (non-fatal) - Non-diabetic - ln(λ) - Age4	0.00001		
MI (non-fatal) - Non-diabetic - ln(λ) - Age5	0.00000		

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
MI (non-fatal) - Non-diabetic - ln(λ) - Sex	1.29600		
MI (non-fatal) - Non-diabetic - ln(λ) - Aspirin treatment	1.15000		
MI (non-fatal) - Non-diabetic - ln(λ) - Insulin treatment	0.00000		
MI (non-fatal) - Non-diabetic - ln(λ) - Statin treatment	3.08600		
MI (non-fatal) - Non-diabetic - ln(λ) - BP drug treatment	2.81600		
MI (non-fatal) - Non-diabetic - ln(λ) - Ex-smoker	-0.82000		
MI (non-fatal) - Non-diabetic - ln(λ) - Smoker	1.07200		
MI (non-fatal) - Non-diabetic - ln(λ) - Constant	-6.72200		
MI (non-fatal) - Non-diabetic - ln(γ) - BMI	0.00000		
MI (non-fatal) - Non-diabetic - ln(γ) - BMI2	0.00000		
MI (non-fatal) - Non-diabetic - ln(γ) - Age	-0.01000		
MI (non-fatal) - Non-diabetic - ln(γ) - Age2	0.00018		
MI (non-fatal) - Non-diabetic - ln(γ) - Sex	-0.08400		
MI (non-fatal) - Non-diabetic - ln(γ) - Aspirin treatment	-0.13900		
MI (non-fatal) - Non-diabetic - ln(γ) - Statin treatment	-0.32800		
MI (non-fatal) - Non-diabetic - ln(γ) - BP drug treatment	-0.24200		
MI (non-fatal) - Non-diabetic - ln(γ) - Ex-smoker	0.08200		
MI (non-fatal) - Non-diabetic - ln(γ) - Smoker	-0.09400		
MI (non-fatal) - Non-diabetic - ln(γ) - Constant	0.90300		
Stroke (non-fatal) - Diabetic - ln(λ) - BMI	0.02000		
Stroke (non-fatal) - Diabetic - ln(λ) - BMI2	0.00000		
Stroke (non-fatal) - Diabetic - ln(λ) - BMI3	0.00000		
Stroke (non-fatal) - Diabetic - ln(λ) - BMI4	0.00000		
Stroke (non-fatal) - Diabetic - ln(λ) - Age	0.05200		
Stroke (non-fatal) - Diabetic - ln(λ) - Age2	0.00000		
Stroke (non-fatal) - Diabetic - ln(λ) - Age3	0.00000		
Stroke (non-fatal) - Diabetic - ln(λ) - Age4	0.00000		
Stroke (non-fatal) - Diabetic - ln(λ) - Age5	0.00000		
Stroke (non-fatal) - Diabetic - ln(λ) - Sex	0.05100		
Stroke (non-fatal) - Diabetic - ln(λ) - Aspirin treatment	1.14400		
Stroke (non-fatal) - Diabetic - ln(λ) - Insulin treatment	0.14600		
Stroke (non-fatal) - Diabetic - ln(λ) - Statin treatment	-0.33100		
Stroke (non-fatal) - Diabetic - ln(λ) - BP drug treatment	0.24600		
Stroke (non-fatal) - Diabetic - ln(λ) - Ex-smoker	-0.40000		

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
Stroke (non-fatal) - Diabetic - ln(λ) - Smoker	0.37600		
Stroke (non-fatal) - Diabetic - ln(λ) - Constant	-9.41000		
Stroke (non-fatal) - Diabetic - ln(γ) - BMI	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - BMI2	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - Age	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - Age2	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - Sex	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - Aspirin treatment	-0.21100		
Stroke (non-fatal) - Diabetic - ln(γ) - Statin treatment	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - BP drug treatment	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - Ex-smoker	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - Smoker	0.00000		
Stroke (non-fatal) - Diabetic - ln(γ) - Constant	0.32500		
Stroke (non-fatal) - Non-diabetic - ln(λ) - BMI	0.03000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - BMI2	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - BMI3	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - BMI4	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Age	0.07300		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Age2	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Age3	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Age4	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Age5	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Sex	0.66400		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Aspirin treatment	1.17600		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Insulin treatment	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Statin treatment	0.60600		
Stroke (non-fatal) - Non-diabetic - ln(λ) - BP drug treatment	0.45000		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Ex-smoker	-1.62700		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Smoker	0.12100		
Stroke (non-fatal) - Non-diabetic - ln(λ) - Constant	-12.14000		
Stroke (non-fatal) - Non-diabetic - ln(γ) - BMI	0.00000		

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
Stroke (non-fatal) - Non-diabetic - ln(γ) - BMI2	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Age	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Age2	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Sex	-0.08000		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Aspirin treatment	-0.18600		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Statin treatment	0.09900		
Stroke (non-fatal) - Non-diabetic - ln(γ) - BP drug treatment	0.00000		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Ex-smoker	0.25500		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Smoker	0.10000		
Stroke (non-fatal) - Non-diabetic - ln(γ) - Constant	0.22000		
T2DM - Non-diabetic - ln(λ) - BMI	0.60700		
T2DM - Non-diabetic - ln(λ) - BMI2	-0.01000		
T2DM - Non-diabetic - ln(λ) - BMI3	0.00006		
T2DM - Non-diabetic - ln(λ) - BMI4	0.00000		
T2DM - Non-diabetic - ln(λ) - Age	0.35700		
T2DM - Non-diabetic - ln(λ) - Age2	-0.00717		
T2DM - Non-diabetic - ln(λ) - Age3	0.00008		
T2DM - Non-diabetic - ln(λ) - Age4	0.00000		
T2DM - Non-diabetic - ln(λ) - Age5	0.00000		
T2DM - Non-diabetic - ln(λ) - Sex	0.79600		
T2DM - Non-diabetic - ln(λ) - Aspirin treatment	-0.19300		
T2DM - Non-diabetic - ln(λ) - Insulin treatment	1.11100		
T2DM - Non-diabetic - ln(λ) - Statin treatment	0.00000		
T2DM - Non-diabetic - ln(λ) - BP drug treatment	-0.38200		
T2DM - Non-diabetic - ln(λ) - Ex-smoker	-0.63700		
T2DM - Non-diabetic - ln(λ) - Smoker	0.28800		
T2DM - Non-diabetic - ln(λ) - Constant	-24.35600		
T2DM - Non-diabetic - ln(γ) - BMI	-0.01100		
T2DM - Non-diabetic - ln(γ) - BMI2	0.00000		
T2DM - Non-diabetic - ln(γ) - Age	-0.01800		
T2DM - Non-diabetic - ln(γ) - Age2	0.00000		
T2DM - Non-diabetic - ln(γ) - Sex	-0.10100		

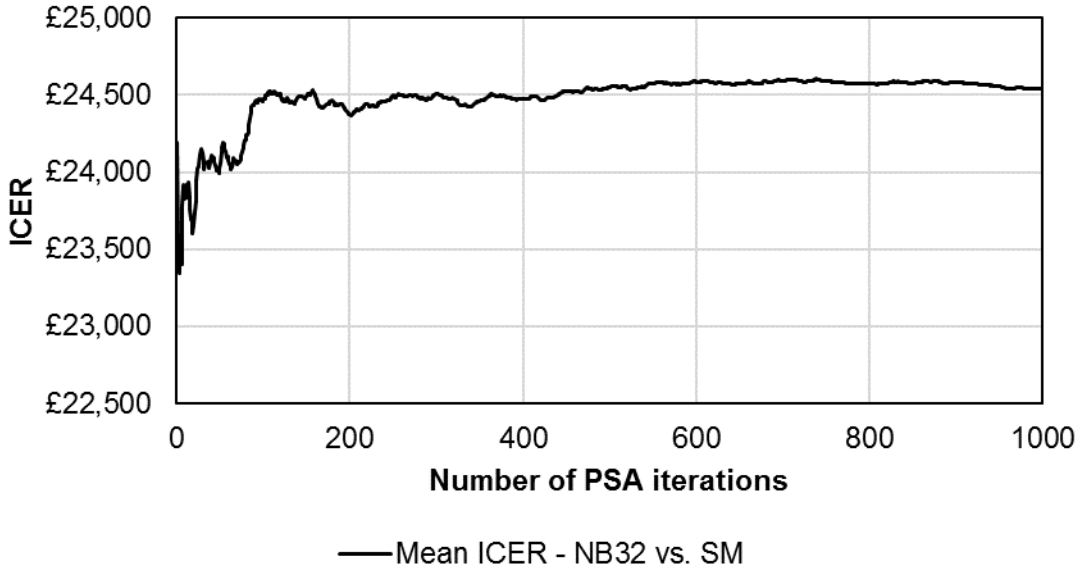
Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
T2DM - Non-diabetic - ln(γ) - Aspirin treatment	0.06600		
T2DM - Non-diabetic - ln(γ) - Statin treatment	0.17700		
T2DM - Non-diabetic - ln(γ) - BP drug treatment	0.14000		
T2DM - Non-diabetic - ln(γ) - Ex-smoker	0.08200		
T2DM - Non-diabetic - ln(γ) - Smoker	-0.03900		
T2DM - Non-diabetic - ln(γ) - Constant	1.32000		
BMI model: Diabetics - Constant	33.17600	Sampled using Normal and multi-variate Normal distributions. SD's for parameters explored as first-order uncertainty taken from Ara *et al.* (discussed in Section 1 of this document).	Section 5.3.4.3
BMI model: Diabetics - Age	0.04000		
BMI model: Diabetics - Sex	-2.06100		
BMI model: Diabetics - Age*Sex	0.00000		
BMI model: Diabetics - SD(Age)	0.33700		
BMI model: Diabetics - SD(Constant)	7.56200		
BMI model: Diabetics - Corr(Age,Constant)	-0.72800		
BMI model: Diabetics - SD(Residual)	1.46600		
BMI model: Non-diabetics - Constant	33.13200		
BMI model: Non-diabetics - Age	0.17500		
BMI model: Non-diabetics - Sex	-2.38100		
BMI model: Non-diabetics - Age*Sex	-0.03000		
BMI model: Non-diabetics - SD(Age)	0.24100		
BMI model: Non-diabetics - SD(Constant)	5.59000		
BMI model: Non-diabetics - Corr(Age,Constant)	0.20400		
BMI model: Non-diabetics - SD(Residual)	1.69200		
Weight loss at 16 weeks (NB32) - Responders	9.40%	SD = 0.032	Section 5.3.3.2
Weight loss at 16 weeks (NB32) - Non-responders	1.89%	SD = 0.024	
Weight loss at 12 weeks (ORL) - Responders	8.27%	Uncertainty sampled via posterior distribution of ITC.	
Weight loss at 12 weeks (ORL) - Non-responders	0.76%		
Weight loss at 12 weeks (SM)	2.27%	SD = 0.036	
Weight loss at 16 weeks (SM)	2.72%	SD = 0.043	
Response at 16 weeks (NB32)	75.65%	SD = 0.012	Section 5.3.3.1
Response at 12 weeks (ORL)	70.54%	Uncertainty sampled via posterior distribution of ITC.	
Weight loss at 56 weeks (NB32) - Primary assessment responders	11.70%	SD = 0.072	Section 5.3.3.2
Weight loss at 56 weeks (NB32) - All patients	8.78%	SD = 0.081	

Variable	Value	Measurement of uncertainty and distribution	Reference to section in submission
Weight loss at 52 weeks (ORL)	10.57%	Uncertainty sampled via posterior distribution of ITC.	
Weight loss at 52 weeks (SM)	3.89%		
Key: ACM, all-cause mortality; BMI, body mass index; BP, blood pressure; CVD, cardiovascular disease; DM, diabetes mellitus; ENT, ear, nose and throat; GP, general practitioner; ITC, indirect treatment comparison; LY, life year; MI, myocardial infarction; NB, naltrexone bupropion; OLS, ordinary least squares; ORL, orlistat; OWSA, one-way sensitivity analysis; QALY, quality-adjusted life year; SD, standard deviation; SE, standard error; SM, standard management.			

3. Results of probabilistic sensitivity analysis

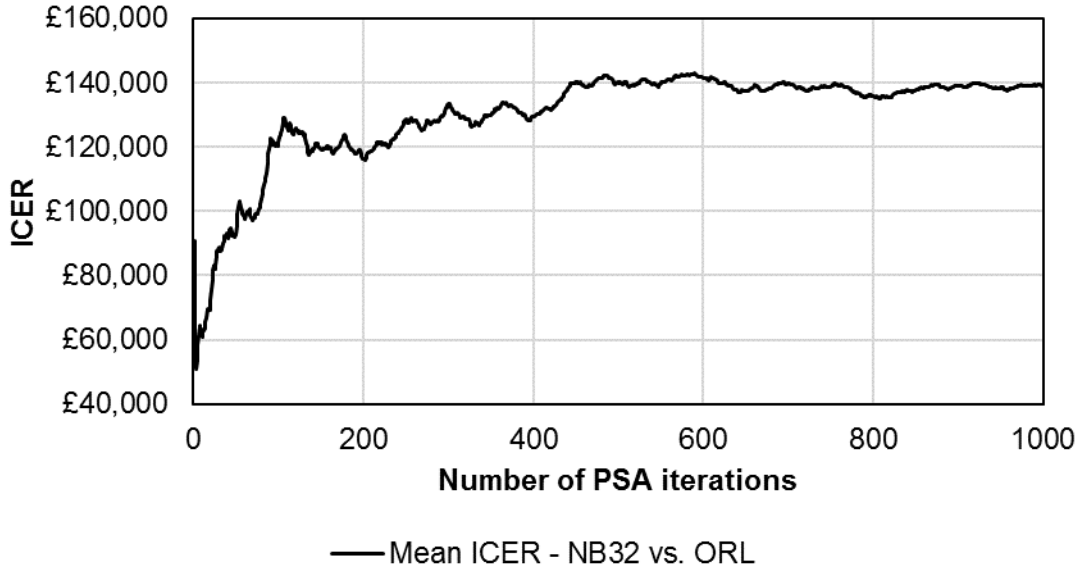
As discussed in Section 1, a sampled patient profile of 15,000 patients was used within each PSA model iteration and 1,000 PSA iterations were simulated for the 15,000 patients, with mean results recorded for each iteration. How the mean pairwise ICERs change as the number of PSA iterations increases is shown in Figure 1 and Figure 2.

Figure 1: Convergence of mean ICER – NB32 vs. SM



Key: ICER, incremental cost-effectiveness ratio; NB32, naltrexone 32mg plus bupropion; PSA, probabilistic sensitivity analysis; SM, standard management.

Figure 2: Convergence of mean ICER – NB32 vs. ORL



Key: ICER, incremental cost-effectiveness ratio; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; PSA, probabilistic sensitivity analysis.

A comparison of the mean probabilistic base case model results with the deterministic base case model results are shown in Table 2. These results are those from the Revised Base Case analysis presented in the 30 May 2017 ACD Response. The probabilistic results are broadly in line with the deterministic results, however due to the sensitivity in the ICER for NB32 versus orlistat, an increase of 0.0017 incremental quality-adjusted life years (QALYs) caused a reduction in the incremental cost-effectiveness ratio (ICER) of approximately £68,656.

Table 2: Comparison of Revised Base Case results: deterministic versus probabilistic

T	Total			Incremental			ICER (QALYs)	
	Costs	LYs[a]	QALYs	Costs	LYs[a]	QALYs	Versus baseline (SM)	Incremental
Deterministic base case model results								
SM	£6,502	33.9109	13.6300					
ORL	£6,802	33.9225	13.6698	£300	0.0116	0.0398	£7,536	£7,536
NB32	£7,531	33.9243	13.6734	£729	0.0018	0.0035	£23,750	£207,274
Probabilistic base case model results								
SM	£6,796	33.8138	13.6159					
ORL	£7,089	33.8246	13.6520	£294	0.0108	0.0361	£8,125	£8,125
NB32	£7,810	33.8268	13.6572	£721	0.0022	0.0052	£24,539	£138,618

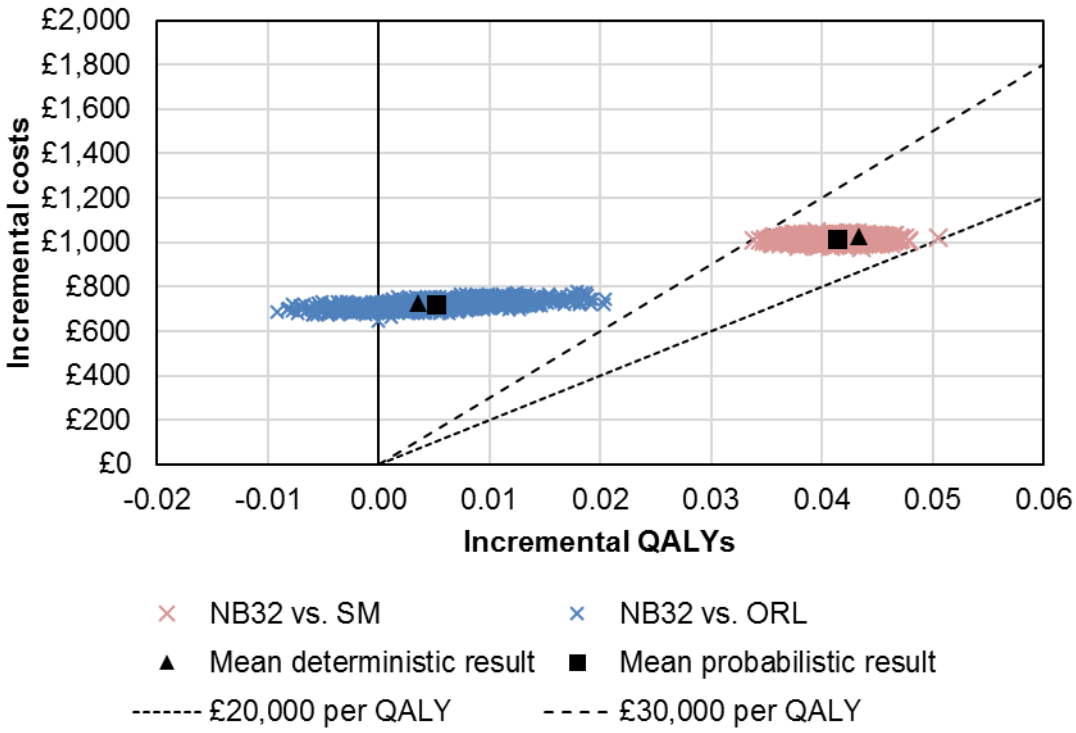
Key: ICER, incremental cost-effectiveness ratio; LY, life year; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management; T, technologies.
Note: [a], LYs are undiscounted, costs and QALYs are discounted.

Figure 3 shows the PSA scatterplot for NB32 versus standard management and NB32 versus orlistat. The scatterplot demonstrates some parameter uncertainty around the mean model result. However, all probabilistic model runs appear to demonstrate results that are not dissimilar to the probabilistic and deterministic mean results.

Importantly, much of the key uncertainty around model results is structural, and based on the key conservative assumptions underpinning the analysis. The uncertainty around results stemming from such uncertainty is not illustrated by probabilistic or deterministic sensitivity analyses.

To have some indication of how important the conservative limitations of the economic analysis are for cost-effectiveness estimates of NB32, the open-access Public Health England (PHE) Weight management economic assessment tool can be considered.[2]

Figure 3: PSA scatterplot



Key: NB32, naltrexone 32mg plus bupropion; PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life year; SM, standard management.

The PHE tool was designed to support public health professionals to make an economic assessment of existing or planned weight management interventions, and considers a cohort approach to provide cost-effectiveness estimates of weight management interventions. The utility data reported in the tool inform utility assumptions in this appraisal. The PHE tool considers the following obesity-related events:

- Stroke
- Coronary heart disease (CHD)
- T2DM
- Colorectal cancer
- Breast cancer

Of these obesity-related events, the only event not directly considered by the submitted model is cancer (CHD in the context of comparing models here is assumed to have a similar impact to MI). Further to this, the PHE model considers excess mortality from these obesity-related diseases. As we have stressed, the economic analysis informing this submission does not.

To gain some estimate of the difference in cost-effectiveness estimates that could be expected if cancer-related QALYs and excess mortality for the diseases were incorporated into Revised Base Case estimates, the PHE tool was downloaded (available at:

http://webarchive.nationalarchives.gov.uk/20170110165405/http://www.noo.org.uk/gsf.php5?f=312732&fv=22235), and user assumptions in the PHE tool were set to match the COR-I study (that is, the proportion of female patients, mean age, mean starting BMI, drop-out rates and average weight loss). Finally, a switch was implemented to omit the impact of the two cancers and excess disease mortality from the estimation of QALYs. The difference in incremental QALYs between NB32 and no treatment (i.e. no weight loss and no cost) was then analysed to establish the impact of the two cancers and excess disease mortality on cost-effectiveness results.

- Inclusion of the two cancers and excess mortality: incremental QALYs = 0.496
- Exclusion of the two cancers and excess mortality: incremental QALYs = 0.494

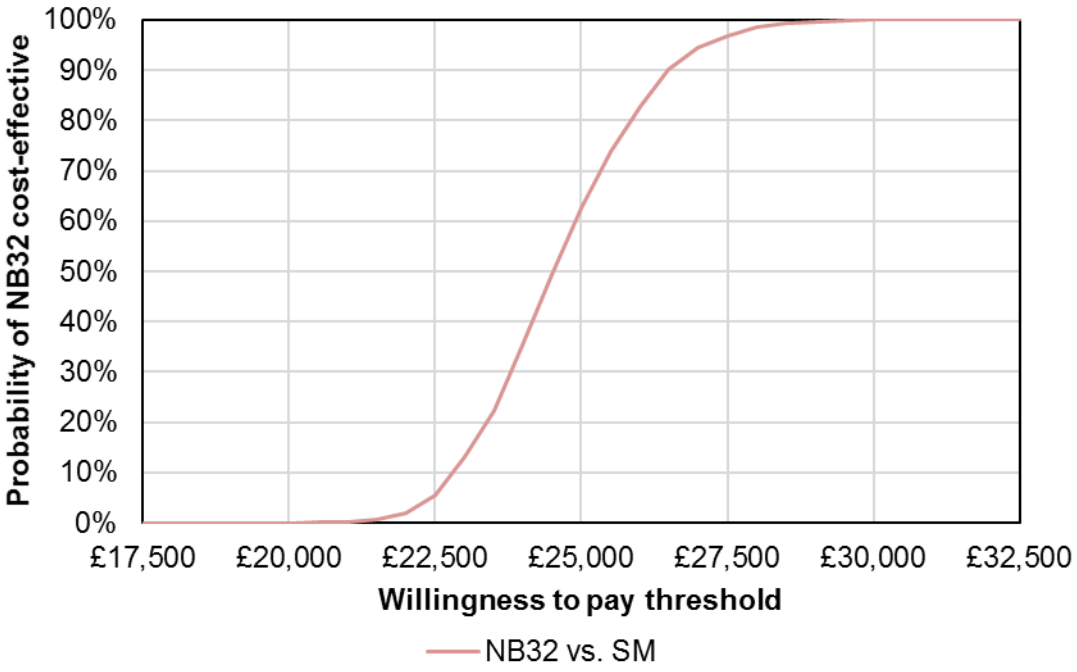
By excluding the two cancers and excess mortality, the PHE tool predicts an under-estimate of around 0.0013 incremental QALYs. Though this may be different to the implications of implementing the cancer risk and excess mortality data into the Revised Base Case model, this gives some indication of the quantitative impact of the conservative approach taken to implementing obesity-related diseases in the submitted *de novo* economic model.

As discussed in our 30 May ACD Response, an additional 0.009 incremental QALYs would reduce the Revised Base Case ICER for NB32 versus standard management to less than £20,000. Therefore, if the inclusion of the remaining 50+ obesity-related health risks and complications listed in the 2015 European Guidelines for Obesity Management in Adults contribute only an additional 0.0077 incremental QALYs, NB32 would be associated with an ICER below £20,000 versus standard management.[3]

Figure 4 shows the cost-effectiveness acceptability curve (CEAC) for NB32 versus standard management. The CEAC shows that for the number of model runs simulated, NB32 is associated with 100% probability of being cost effective versus standard management at a willingness to pay (WTP) threshold of £30,000 per QALY gained.

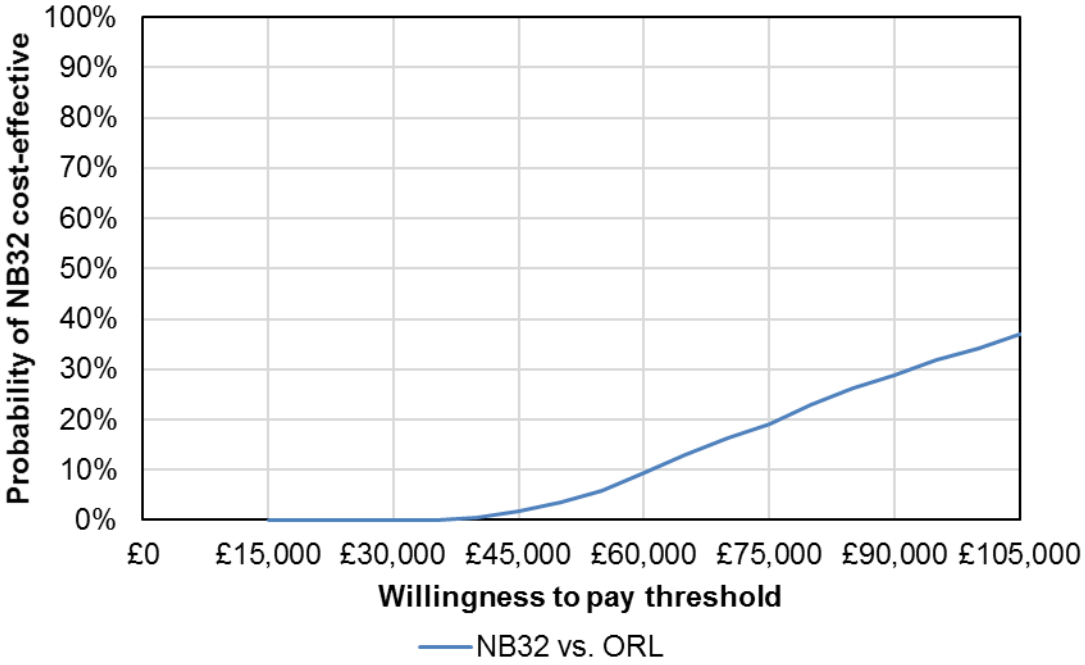
Figure 5 shows the CEAC for NB32 versus orlistat up to a maximum WTP of £105,000 per QALY gained. As previously discussed, much of the key uncertainty around model results is structural or methodological, and based in the key conservative assumptions underpinning the analysis. It is highly plausible that the estimated probability that NB32 is preferable to orlistat would be greater if just some of the downstream health and cost benefits of weight loss for obesity-related health events not currently informing the model could be captured.

Figure 4: CEAC – NB32 vs. SM



Key: CEAC, cost-effectiveness acceptability curve; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life year; SM, standard management.

Figure 5: CEAC – NB32 vs. ORL



Key: CEAC, cost-effectiveness acceptability curve; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life year; SM, standard management.

References

1. Greenwood M. *A Report on the Natural Duration of Cancer*. H.M. Stationery Office, 1926.
2. Copley V. *User Guide: Weight Management Economic Assessment Tool Version 2*. Oxford: Public Health England, Obesity Risk Factors Intelligence, 2016.
3. Yumuk V, Tsigos C, Fried M, et al. European Guidelines for Obesity Management in Adults. *Obes Facts*. 2015; 8(6):402-24.

Single Technology Appraisal (STA): Naltrexone-bupropion (prolonged-release) for managing overweight and obesity

Response to Appraisal consultation document

1. **Has all of the relevant evidence been taken into account?**

I am not aware of any relevant studies other than those already included.

2. **Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?**

The summaries appear to be appropriate. The evidence demonstrates naltrexone-bupropion to be clinically effective in comparison to placebo, and indirect comparisons suggest, on the whole, similar clinical efficacy between naltrexone-bupropion and orlistat.

With regard to assessment of cost effectiveness, it is very disappointing that it has not been possible to make a sufficiently reliable assessment of cost effectiveness, as it is on this basis that the conclusion has been reached that naltrexone-bupropion cannot be recommended as an option for managing overweight and obesity within the NHS. However, the basis on which this conclusion has been reached is clearly explained.

3. **Are the provisional recommendations sound and a suitable basis for guidance to the NHS?**

There is a clear clinical need for novel pharmacological approaches to treatment of overweight and obesity, and naltrexone-bupropion is a novel treatment that could be well placed as part of an integrated weight management pathway. It must always be emphasised that it is important for weight management services to be available within a cohesive four tier pathway, including specialist non-surgical specialist weight management services at Tier 3 and bariatric surgery at Tier 4. Against this background, it is deeply disappointing that naltrexone-bupropion cannot currently be recommended. However, in the absence of a reliable estimate of cost effectiveness, this recommendation does appear to be appropriate.

[redacted]

On Behalf of the Royal College of Pathologists

29th May 2017

Royal College of Physicians
11 St Andrews Place
Regent's Park
London NW1 4LE
Tel: +44 (0)20 3075 1560

www.rcplondon.ac.uk

From The Registrar
[redacted]

National Institute for Health and Care Excellence
10 Spring Gardens
London
SW1A 2BU
TACommA@nice.org.uk

26 May 2017

Dear Sir or Madam

Re: Naltrexone–bupropion (prolonged-release) for overweight and obesity

The Royal College of Physicians (RCP) plays a leading role in the delivery of high quality patient care by setting standards of medical practice and promoting clinical excellence. We provide physicians in the United Kingdom and overseas with education, training and support throughout their careers. As an independent body representing over 33,000 Fellows and Members worldwide, we advise and work with government, the public, patients and other professions to improve health and healthcare.

The RCP is grateful for the opportunity to respond to the above consultation. We have liaised with our Nutrition Committee and would like to make the following comments.

- There is significant unmet need for pharmacotherapy for obesity; there is a strong patient voice that is asking for this.
- The only currently available treatment (orlistat) that was used for comparison is not tolerated by a very high proportion of patients.
- The valid comparator should be lifestyle intervention, not orlistat.
- One possibility would be to support use in patients who do not respond to or are intolerant of orlistat.
- It is difficult to get into the discussion about the economic data as it needs more work.

Yours faithfully

[redacted]
Registrar

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Naltrexone-bupropion for managing overweight and obesity
ADDENDUM

Produced by

Kleijnen Systematic Reviews (KSR) Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors

Debra Fayter, Systematic Reviewer, KSR Ltd, UK
Bram Ramaekers, Health Economist, Maastricht UMC
Sabine Grimm, Health Economist, Maastricht UMC
Nigel Armstrong, Health Economist, KSR Ltd
Caro Noake, Information Specialist, KSR Ltd
Ching-Yun Wei, Health Economist, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Sofia Carrera, Systematic Reviewer, KSR Ltd
Rob Riemsma, Reviews Manager, KSR Ltd
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to

Rob Riemsma, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, UK
YO19 6FD

Date completed

06/06/2017

1. Critique of the company's new submission in response to the ACD

1.1 Overview of the new submission

In response to the Appraisal Consultation Document (ACD), the company submitted a new implementation of its economic model and implemented a number of the analyses and amendments requested by the committee, in what is referred to as the company's revised base-case. The new implementation of the model is efficient and appears well done and well documented but could not be validated by the Evidence Review Group (ERG) due to the extremely tight time constraints.

In the revised base-case, the new incremental cost effectiveness ratio (ICER) of NB32 when compared with orlistat was £207,274 per quality-adjusted life year (QALY) gained, and when compared with standard management (SM) £23,750 per QALY gained. The large increase in the ICERs was predominantly caused by the addition of an annual updating event, as previously suggested by the ERG, to more accurately reflect changes in Body Mass Index (BMI) and therefore changes to the risk of patients of having a cardiovascular (CV) event or experiencing the onset of Type 2 Diabetes Mellitus (T2DM) as well as health-related quality of life (HRQoL) (causing an increase in the ICER of NB32 compared with orlistat of approximately £110,000 per QALY gained).

The company presented new arguments to be considered in decision-making. The following paragraphs describe and critique the company's new implementation of the model, as well as these other arguments presented by the company.

1.2 Model implementation

The new implementation of the model is significantly more efficient (with model run times of 11 minutes for the deterministic analysis on the ERG's computers, probabilistic sensitivity analysis (PSA) was estimated to take 7 days for 1,000 PSA runs on one single computer by the company) and draws on 15,000 patient samples. The ERG was able to deterministically run the model a few times and obtained ICERs fairly close to the ones obtained by the company when new random numbers and patient profiles were used, although the NB32 versus orlistat ICER differed by up to £2,000 per QALY gained. This suggests that using 15,000 patient samples does not achieve complete stability of results, owing to the very small QALY gains provided by NB32 over orlistat (0.0035 incremental QALYs in the company's revised base-case, company's model run), but could be deemed sufficient for decision-making in this instance given that the resulting ICERs do not cross the likely maximum acceptable ICER.

The company undertook validation of this new implementation, resolved a few technical errors that they identified in the previous version and provided a comparison between the old and new implementation of the model in terms of the ICER and other outcomes, such as life years, QALYs, drug costs, standard management and condition management costs, adverse event costs, and death costs. These results give confidence that the new implementation of the model is indeed doing what it is supposed to. However, it is important to note that the ERG was not given the time to perform appropriate, or any, validation of this newly submitted model.

While trying to run the model, the ERG noticed that the macro for generating new patient profiles was not working and was still limited to generating 1,000 patient profiles. The ERG fixed the macro and enabled it to generate 15,000 patient profiles to ensure that patient heterogeneity was fully reflected in the analysis using 15,000 patient samples. The ERG furthermore noticed an error in the generation of patient profiles resulting in patients with a negative age in an extremely small number of cases, due to the use of a normal distribution in sampling the age.

The company provided the PSA with a slight delay. Given the short timelines, the ERG was unable to validate the implementation of the PSA. The ERG notes that the company attempted to reflect uncertainty about time to treatment discontinuation (TTD), the BMI trajectory model and the time to obesity-related events, as was requested by the committee. The company ran a probabilistic analysis using 1,000 PSA runs and 15,000 patient samples in each PSA run. Based on the diagnostic exercise provided by the company, 1,000 PSA runs appear to give a good indication of probabilistic results (even though more PSA runs would be desirable). The ERG notes that, given the extremely small magnitude of the incremental QALY gains, it would be preferable to report diagnostics in terms of incremental costs and QALYs rather than the ICER. Furthermore, probabilistic results vastly differ from deterministic results (NB32 versus orlistat probabilistic ICER of £138,618 per QALY gained, with an incremental QALY gain of 0.0052, compared with the deterministic ICER of £207,274 per QALY gained, with an incremental QALY gain of 0.0035). It is not unusual for DES models to be non-linear and to therefore exhibit differences in probabilistic versus deterministic results. Especially given the volatility of the ICER due to the small magnitude of incremental QALYs, this large difference in probabilistic versus deterministic results may be an artefact of this non-linearity. However, the ERG would have liked to see this issue explored, to make sure that the possibility of a systematic cause for any over-estimation of incremental QALYs can be excluded.

The large difference in ICERs when comparing the company's original and the company's revised base-case stems to a large part from the addition of an annual updating event. This annual updating event is essential in capturing the change in BMI of each modelled patient at every year. In the original company's base-case, the lack of an updating event meant that the BMI was only re-calculated every time an event occurred, which on average resulted in BMI and risk equations being updated only once every 17 years. The addition of the updating event means that the BMI trajectory is appropriately reflected every year, which in turn leads to a more accurate estimation of each patient's HRQoL and risk of CV events, onset of T2DM and time to death.

1.3 The appropriate comparator for this decision problem / scenario after orlistat failure

In section 1.1 and 1.2 of the company response to the ACD, the company suggests that "the committee may want to consider NB32 as an alternative to SM alone in patients who fail orlistat treatment." That would make standard management without pharmacological treatment the sole relevant comparator for patients who have failed previous orlistat treatment.

This is not a population defined in the scope and there are several problems with this comparison. First of all, there is no evidence presented for this population. As explained by the company, orlistat use prior to the 4 weeks before entering the COR trials was not documented. Therefore, it is unclear what the size of

this population is and what the relative effectiveness of NB32 is when compared to standard management in this population. The company argues that orlistat and NB32 have different mechanisms of action and that previous orlistat treatment is not expected to affect NB32 treatment effectiveness. However, no evidence for this assumption is presented. Moreover, the effectiveness of NB32 relies not only on the drug itself but also on standard management and the patient's behaviour. Since NB32 and orlistat are only parts of a multi-component treatment, it would be misleading to assume independence in treatment effectiveness and this might well result in biased model outcomes. A second problem is that it might be difficult to define orlistat failure; is it patients who are intolerant to orlistat, or also those who are non-responders?

1.4 The suitability of the indirect comparison of NB32 and orlistat

In section 1.3 of the company response to the ACD the company present several limitations when comparing the clinical benefits of NB32 to that of orlistat. While we agree that there are limitations to any indirect comparison, we fail to see why these limitations would necessarily favour either NB32 or orlistat.

The company state that orlistat studies do not adhere to the baseline observation carried forward imputation method used for NB32 trial data, and that this biases the ITC between orlistat and NB32 in favour of orlistat. However, the company presents no evidence for this statement. Likewise, it could be argued that some differences favour NB32. For instance, the four NB32 studies are more recent (2010, 2011 and 2013(2x)) than the eight orlistat studies (2000, 2001, 2002 (2x), 2005, 2006, 2010 and 2012) mentioned in section 1.3 of the company response to the ACD. It could be argued that improved dietary advice and behavioural interventions, in addition to a better awareness of the risks of obesity over time, make more current standard management more beneficial than that of a decade earlier. However, there is no evidence for any of these speculations.

1.5 The issue of the model potentially underestimating the benefit of NB32

The company argued that their model under-estimated the benefits that can be obtained with NB32, as the costs and HRQoL implications of risks for other conditions that could be influenced by weight reduction in obese and overweight patients were not incorporated. The company, in their provision of the updated probabilistic sensitivity analysis results document, give an indication of the impact of including the obesity-related risk of cancer in the analysis. This was done by using the Public Health England (PHE) weight management economic assessment tool, which includes colorectal and breast cancer as potential obesity-related events. Patient characteristics were set to match the COR-I study and the incremental QALY gain of NB32 patients compared with no treatment was estimated with and without the inclusion of the two cancers and excess mortality. The company uses the resulting 0.0013 incremental QALYs as an indication for how much the company's revised base-case model may under-estimate the benefits of NB32, although the company admits that this may be different if implemented in their revised base-case model. It is the ERG's opinion that the model may indeed be an over-simplification of the actual experience of obese and overweight patients. However, the ERG would like to see these costs and HRQoL implications of other conditions and mortality incorporated in the model for the SM, orlistat and NB32 arms, to be able to inspect the true impact on model outcomes. If this is done, evidence derived from a systematic review should be used to inform the additions to the model.

2. Critique of the company's economic analyses in response to the ACD

2.1 Overview of company's changes compared with previous ERG critique

Table 1 provides an overview of all issues previously highlighted by the ERG that were addressed by the company (part A), those that are potentially unresolved (part B) and new analyses provided by the company (part C).

The company made a number of changes to their original base-case in their company's revised base-case that were requested by the committee (Table 1, part A). These changes include the increase of patient samples run, implementation of annual updates in the model, the use of patient baseline characteristics in line with the COR trial programme, the use of the intention-to-treat (ITT) population from COR-I and COR-DM trials instead of the modified ITT (mITT) analysis from the entire COR trial programme, the removal of linear scaling from orlistat time to treatment discontinuation (TTD) estimates, the incorporation of bariatric surgery as a consequence for patients, and the removal of one GP visit for patients receiving SM. The company furthermore improved their implementation of the PSA. A more detailed critique is provided in Section **Error! Reference source not found.**.

The company made changes that were not requested by the committee (Table 1, part C), notably:

(1) The use of TTD from COR-I and COR-DM only instead of the entire COR trial programme, which is in line with other recommendations regarding the use of evidence. The ERG agrees with this approach.
(2) The addition of a scenario in which NB32 is only used after orlistat failure and compared with SM only. This scenario, in the ERG's view, has to be interpreted with extreme caution. This is because there is no evidence to inform such a scenario (please see Section 1.3 for more detail).
(3) Furthermore, the company fixed a few modelling errors that were identified in re-programming the model and which drove the ICER up. The ERG was unable to verify these changes but they appear to have face validity.

There are a number of issues that are potentially unresolved (Table 1, part B). The ones that the ERG considers to be cause for concern are detailed in the following:

(1) Model structure: the model remains restricted to only two cardiovascular events, and continues to omit other potential comorbidities and diseases.
(2) Comparators: the model does not include behaviour interventions.
(3) Model assumption: three years after treatment discontinuation patients regain their weight to the predicted instead of the baseline BMI. As was stated before, this is a non-conservative assumption (ICER of NB32 versus orlistat is smaller with this assumption than with the alternative) and is not in line with Ara et al.'s model.[1] The company provides an example that illustrates why their assumption may exhibit face validity. The ERG agrees with the company's conclusion that it is an implausible artefact of Ara et al.'s[1] baseline assumption that a non-responder would achieve a life-long benefit. However, the ERG considers it as equally implausible that a responder only experiences a benefit for the time that they continue treatment (which in the model is on average only 13 months) and for the three years afterwards, at which point they have reverted back to their originally predicted BMI trajectory without any benefit for their future life. However, there is uncertainty about this and in the absence of further evidence or clinical opinion to inform this

issue for both responders and non-responders separately, the ERG prefers the more conservative course of action, which is a weight regain to the baseline BMI.

(4) TTD: the TTD after 56 weeks continues to be likely an under-estimate because it is derived from the NB-CVOT study, in which patients were at increased risk of adverse cardiovascular outcomes.

(5) Modelling of adverse events (AE): AE rates are still only informed by COR-I, and not COR-DM. Equivalent costs for AEs are still used for NB32 and orlistat.

(6) Effectiveness measure: the company did not change their approach of assuming that the absolute mean difference (MD) in change in weight from baseline at 52/56 weeks would hold true for change in weight from baseline at 12/16 weeks to the relative approach suggested by the ERG. The justification provided by the company that the ERG's approach resulted in counter-intuitive results due the small denominator was deemed appropriate. However, there are other ways of adjusting the absolute MD to obtain a more appropriate relative MD by standardising the MD by the average weight at the respective assessment.

Table 1: Issues in the economic model and how they were addressed

Issue	Bias introduced[a]	ERG analyses (analysis number in section 5.3)	*Addressed by company in response to ACD/ERG comment*
A) Changes implemented suggested by the committee / ERG			
• Weight regain is not implemented linearly in the economic model.	+	Base-case (1)	*This has now been addressed through the annual updating event.*
• Pooling from all COR studies is inappropriate because: 1. BMOD uses different intensity of treatment- accompanying management; 2. COR-II data are only available up to 28 weeks.	+	Base-case (2)	*This has been addressed based on the ERG's correction.*
• Baseline BMI is vastly underestimated in the economic model, hence overestimating utility and time to T2DM, cardiovascular events and death.	+	Base-case (4)	*This has been addressed based on the ERG's correction.*
• The proportions of current smokers, patients receiving anti-hypertensive medication and/or statins are most likely underestimated.	+	Base-case (5)	*This has been addressed based on the ERG's correction.*
• Counter-intuitive patient profiles are generated as correlations between patient characteristics are not incorporated.	+/-	Base-case (5)	*This has been addressed based on the ERG's correction.*
• An unnecessary GP visit, related to response assessment, is incorporated for standard management.	+	Base-case (6)	*This has been addressed based on the ERG's correction.*
• TTD data for orlistat were obtained by linearly scaling the 16 weeks TTD curve for NB32 to fit into the 12 weeks. The company did not provide alternative analysis upon request.	+/-	Base-case (8)	*This has been addressed based on the ERG's correction.*
• Bariatric surgery as a subsequent treatment option was not implemented.	+/-		*This has now been addressed.*
• The number of simulated patients (1,000) is too low to provide stable results; the ICER varies substantially with each model run.	+/-		*This has now been addressed.*

Issue	Bias introduced[a]	ERG analyses (analysis number in section 5.3)	*Addressed by company in response to ACD/ERG comment*
• The lack of an updating event overestimates utility and time to T2DM, cardiovascular events and death. Moreover, implicitly assuming a stable BMI for on average 17 years hampers the face validity of the model. • The PSA does not include TTD, time to obesity-related events and BMI trajectory and uses too small a number of PSA runs and patient samples in each individual PSA run.	+		*An updating event (annual) was added to the analysis.* *This has now been addressed.*
B) Potentially unresolved issues			
• Behavioural interventions and re-treatment are not implemented.	+/-		*Behavioural interventions and re-treatment are still not implemented. Evidence is lacking for re-treatment .*
• Weight regain assumptions deviated from those in Ara et al.[1] in that the company modelled weight regain towards the predicted BMI instead of the baseline BMI.	+	Base-case (7)	*The company maintained weight regain towards predicted BMI instead of baseline BMI.*
• The model structure is restricted to only having two cardiovascular events. Experiencing a stroke after two MI's might have an impact on the outcomes and costs.	-		*This has not been addressed.*
• An assumption is made that weight loss is equivalent for NB32 and orlistat at different times (16 weeks and 12 weeks, respectively).	+/-		*This assumption is still being made.*
• The mean change in weight for orlistat at primary assessment was derived using the mean difference in treatment effect at secondary assessment (for NB32 versus orlistat) applied to NB32 mean change in weight at primary assessment	+	Base-case (3)	*This has not been addressed by the company and appropriate justification was provided for not accepting the ERG's approach. However, other approaches to reflect a relative MD are possible.*

Issue	Bias introduced[a]	ERG analyses (analysis number in section 5.3)	*Addressed by company in response to ACD/ERG comment*
• TTD (after 56 weeks) is under-estimated because it was derived from a more severe patient population (from NB-CVOT study) and it was assumed that all patients discontinued after the trial period had ended.	+		*This has not been addressed.*
• It is questionable whether the results of the economic analyses are representative for patients with a history of angina and/or diabetes other than T2DM.	+/-		*This has not been addressed.*
• AE-related utility decrements were not included.	+		*This has not been addressed but was shown not to be influential.*
• Only the COR-I trial was used to inform AE rates; the COR-DM trial could have been used to obtain T2DM specific AE rates.	+/-		*This has not been addressed.*
• It is questionable whether the assumption of equivalent AE costs for NB32 and orlistat is conservative.	+/-		*This has not been addressed.*
• Use of PHE weight management economic assessment tool for derivation of utilities may not be appropriate.	+/-		*This has not been addressed.*
• Assuming only a single GP visit for each adverse event without plausible justification.	+		*This has not been addressed.*
• NB32 drug wastage was not considered in the model	+		*This has not been addressed.*
C) Other additions / analyses			
• Use TTD from COR-I and COR-DM only instead of entire COR programme	+/-		*Deemed appropriate by ERG.*
• Added scenario of NB32 vs SM after treatment failure with orlistat	+/-		*Has to be interpreted with extreme caution.*

Abbreviations: NA, not applicable

[a]Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' in indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator.

2.2 Company's revised base-case results

The company's revised base-case results differed significantly from their original base-case results (by approximately £162,000 per QALY gained in the comparison with orlistat) (see Table 2). The company provided a step-by-step overview to show each change made to the model and its impact on the ICER. This showed that the by far most influential change (change of approximately £110,000 per QALY gained in the comparison with orlistat) was to implement an annual updating event to reflect the progressive increase in BMI that overweight and obese patients typically experience, and to appropriately reflect the change in risk of CV events and onset of diabetes as well as HRQoL associated with BMI. This highlights the importance of an annual (or even more frequent) updating event in Discrete Event Simulation (DES) models in which there is a progressive outcome measure, such as BMI in this case.

The second most influential change was a combined one of fixing various technical errors and that accounted for a change of approximately £50,000 per QALY gained in the comparison with orlistat. Using TTD data from COR-I and COR-DM increased the ICER by approximately £23,000 per QALY gained in the comparison with orlistat. Some other minor changes to the model reduced the ICER by approximately £16,000 per QALY gained in the comparison with orlistat. The addition of bariatric surgery barely influenced model outcomes.

Table 2: Company's revised base-case model results

Technologies	Total			Incremental			ICER	
	Costs	LYs	QALYs	Costs	LYs	QALYs	Versus baseline [SM]	Incremental
Incremental								
SM	£6,502	33.9109	13.6300					
ORL	£6,802	33.9225	13.6698	£300	0.0116	0.0398	£7,536	£7,536
NB32	£7,531	33.9243	13.6734	£729	0.0018	0.0035	£23,750	£207,274
Pairwise								
SM	£6,502	33.9109	13.6300					
NB32	£7,531	33.9243	13.6734	£1,029	0.0134	0.0433	£23,750	
ORL	£6,802	33.9225	13.6698					
NB32	£7,531	33.9243	13.6734	£729	0.0018	0.0035	£207,274	

Key: ICER, incremental cost-effectiveness ratio; LYs, life years; NB32, naltrexone bupropion; ORL, orlistat; QALYs, quality-adjusted life years; TTD, time to treatment discontinuation.

2.3 Company's subgroup analysis results

As in the previous submission, model results were more favourable for NB32 in the non-diabetic subgroup (ICER versus orlistat of £57,899 per QALY gained) than in the diabetic subgroup (NB32 dominated by orlistat).

2.4 Company's scenario analysis results

The company's scenario analysis exploring cost effectiveness of NB32 compared with SM after treatment failure with orlistat resulted in an ICER versus SM of £23,324 per QALY gained. It is important to note that this is based on the assumption that treatment effectiveness of NB32 is independent of prior treatment (success or failure), and that there is no evidence to support this assumption. In contrast, it could be argued that independence cannot hold, as was detailed in Section 1.3. The ERG therefore considers there to be significant uncertainty associated with the estimate of the ICER. The small incremental QALY gains this ICER is based on imply that even small changes in the effectiveness estimate may result in large differences in the ICER.

References:

[1] Ara R, Blake L, Gray L, Hernandez M, Crowther M, Dunkley A, et al. What is the clinical effectiveness and cost-effectiveness of using drugs in treating obese patients in primary care? A systematic review. *Health Technol Assess* 2012;16(5):iii-xiv, 1-195.

Comments on the ACD Received from the Public through the NICE Website

Name	[redacted]
Role	NHS Professional
Other role	[redacted]
Organisation	University of Liverpool and University Hospital Aintree
Location	England
Conflict	I have worked alongside a wide variety of Pharma companies and have received consultancy fees and research funding from a number of companies. I am aware of the trial data relating to all these drugs and feel they all must form a component of our strategy to manage obesity in the UK
Notes	
Comments on individual sections of the ACD:	
Section 1 (Appraisal Committee's preliminary recommendations)	General comment: Working as a metabolic physician with a specialist interest in Obesity and type 2 diabetes as well as being an active obesity clinical researcher, the lack of therapeutic options in obesity in the UK is startling. Lifestyle intervention is undoubtedly effective but in reality the proportion in whom it is effective is very limited and of limited magnitude. We have a significant unmet need with up to 4 drugs available in countries like the US. The use of orlisatat is incredibly limited by GPs or specialists and is poorly tolerated. For this reason I believe this drug should be made available as a therapeutic option. I would also suggest comparing the response against orlistat is not valid due to the very small number of patients on it. I would suggest lifestyle is the relevant comparator. Application of appropriate stopping rules ensure this drug would be continued in those who derive clinical benefit and I would urge NICE to endorse this drug so that we may impact upon the obesity epidemic.
Section 2 (The technology)	
Section 3 (The manufacturer's submission)	
Section 4 (Consideration of the evidence)	
Section 5 (Implementation)	
Section 6 (Related NICE guidance)	
Section 7 (Proposed date of review of guidance)	