NICE National Institute for Health and Care Excellence

Single Technology Appraisal

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

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

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

Contents:

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

Premeeting briefing
Naltrexone-bupropion (prolonged-release) for managing overweight and obesity (ID757)

This slide set is the pre-meeting briefing for this appraisal. It has been prepared by the technical team with input from the committee lead team and the committee chair. It is sent to the appraisal committee before the committee meeting as part of the committee papers. It summarises:

- the key evidence and views submitted by the company, the consultees and their nominated clinical experts and patient experts and
- the Evidence Review Group (ERG) report.

It highlights key issues for discussion at the first appraisal committee meeting and should be read with the full supporting documents for this appraisal.

Please note that this document includes information from the ERG before the company has checked the ERG report for factual inaccuracies.

The lead team may use, or amend, some of these slides for their presentation at the Committee meeting.

Key issues: clinical effectiveness

1. **Positioning**
   - What is the expected positioning of naltrexone-bupropion in the treatment pathway: alternative to orlistat vs 2nd line to orlistat? What are the relevant comparators? Are different types of behaviour modification, such as more intensive forms, relevant?
2. **Population**
   - should effectiveness be considered in a mixed population (overweight and obese) with and without Type 2 Diabetes Mellitus (T2DM)?
3. **Effectiveness of naltrexone-bupropion (NB32) vs placebo**
   - What is the appropriate analysis: Intention to treat (ITT) or modified ITT (mITT)? Implications of large drop out rate, and how to deal with this analytically: Last observation carried forward, Baseline observation carried forward, Weight regain imputation? Should the COR trials be pooled?
4. **Effectiveness of NB32 vs orlistat**
   - Which trials should be used in the indirect treatment comparison of NB32 vs orlistat?
5. **Generalisability to NHS**
   - Is standard management in the COR trials generalisable? Is the patient population in the trials generalisable?

2

Overweight and obesity

- Chronic condition characterised by increased body fat – people are at an increased risk of developing cardiovascular disease (CVD), Type-2 diabetes mellitus (T2DM), hypertension, dyslipidaemia and atherosclerosis
- Body Mass Index (BMI) is the most common method for measuring obesity:
  - 25 kg/m^2 to 29.9 kg/m^2: overweight
  - 30 kg/m2 to 34.9 kg/m^2: obese I
  - 35 kg/m2 to 39.9 kg/m^2: obese II
  - 40kg/m2 or more: obese III
- Prevalence
  - In England, 24% of adults are obese and a further 36% are overweight
  - 7/10 are class 1 obese (BMI of 30 – 34.9), and 1/10 morbidly obese (BMI of 40 or more)
  - Expected prevalence of obesity in 2050 - 60% of adult men and 50% of adult women

3

Source: see section 3 of the company's submission.

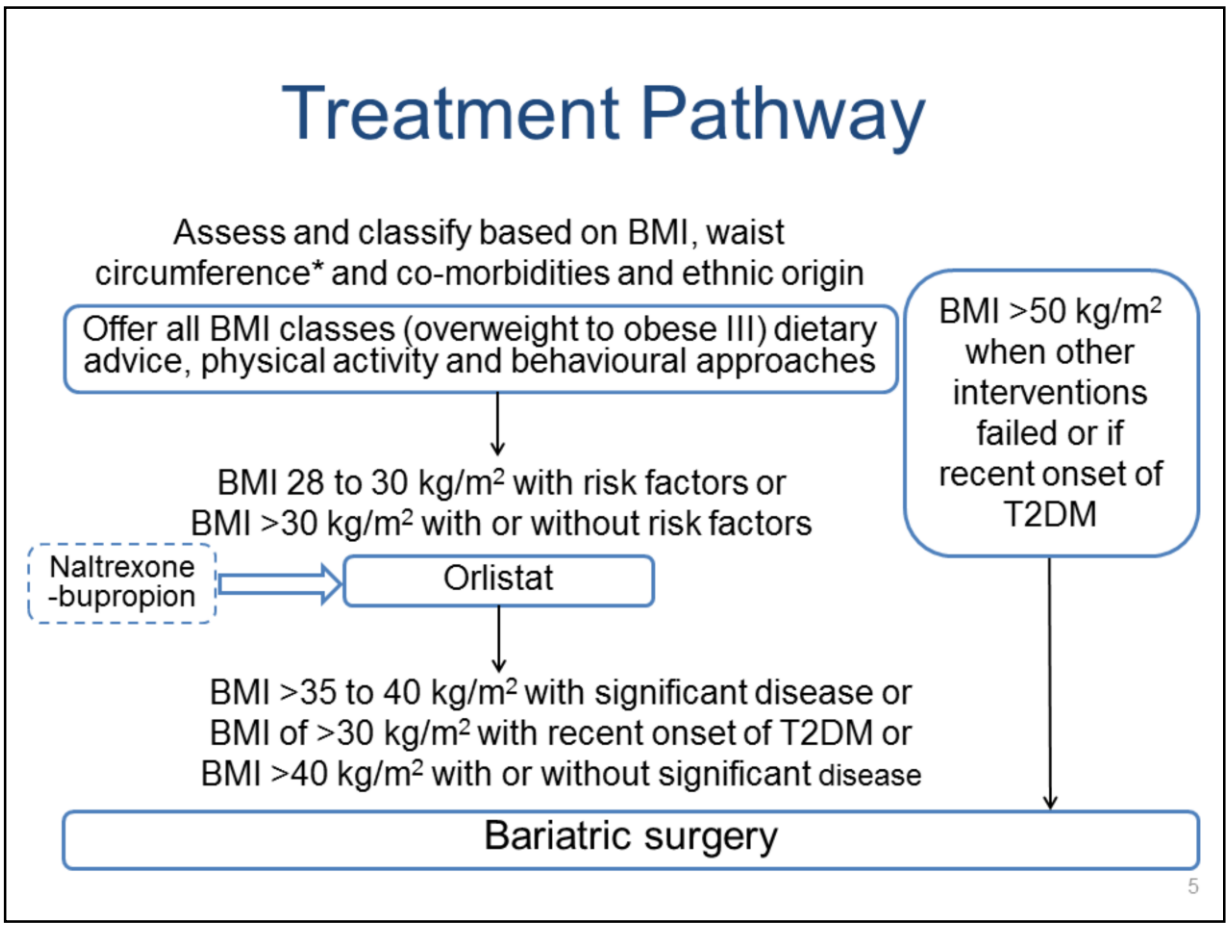
Naltrexone-bupropion (NB32)

(Naltrexone 32mg plus bupropion 360mg prolonged-release)

UK marketing authorisation	'Adjunct to a reduced-calorie diet and increased physical activity, for the management of weight in adult patients (≥18 years) with an initial BMI of • ≥ 30 kg/m^2 (obese), or • ≥ 27 kg/m^2 to < 30 kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities (e.g. type 2 diabetes, dyslipidaemia, or controlled hypertension) Treatment should be discontinued after 16 weeks if patients have not lost at least 5% of initial body weight'
Class of drug	Naltrexone is an opioid receptor antagonist and bupropion is a dopamine and noradrenaline reuptake inhibitor. Exact neurochemical effect is unknown but is thought to stimulate pro-opiomelanocortin neuronal firing and modulate food cravings through an effect on the reward pathways of the brain.
Administration and dosage	Administered orally in a prolonged-release tablet. Dose is escalated over a 4-week period to a total dose of 32 mg naltrexone and 360 mg bupropion: Week 1: one tablet in morning; Week 2: one tablet morning & evening; Week 3: two tablets in morning & one in evening; From week 4: two tablets morning & evening
Cost	Acquisition cost (excl. VAT) £73.00 per pack of 112 tablets Predicted lifetime cost £995 (Source: company's submission)

4

Source: see section 2 of the company's submissions.



Treatment pathway as recommended in 'Obesity: the prevention, identification, assessment and management of obesity in adults and children' (clinical guideline 189).

*For men, waist circumference of less than 94 cm is low, 94–102 cm is high and more than 102 cm is very high, for women, waist circumference of less than 80 cm is low, 80–88 cm is high and more than 88 cm is very high.

Current management

Weight management in England is based on well-defined tier services for which a person is grouped into and receives care based on an assessment of BMI, waist circumference and the presence of comorbidities. Tier 1 comprises universal services such as health promotion, Tier 2 covers lifestyle interventions, Tier 3 covers services, and Tier 4 covers bariatric surgery.

specialist weight management

Orlistat is the only approved drug treatment available in the UK (covered in tier 3 services). NICE clinical guideline 189 'Obesity: identification, assessment and management' recommends that orlistat should only be considered after dietary, physical activity and behavioural approaches have been started and evaluated. It recommends orlistat for the management of obesity in people with a BMI of 30 kg/m^2 or more, and in people with a BMI of 28 kg/m^2 or more and significant comorbidities. If dietary and lifestyle advice, behaviour modification and drug treatments are unsuccessful, the NICE clinical guideline recommends bariatric surgery for people with: a BMI of 40 kg/m^2 or more; a BMI of between 35 kg/m^2 and 40 kg/m^2 with significant comorbidities, a BMI between 30 kg/m^2 and < 35 kg/m^2 and with recent-onset of type

2 diabetes (surgery can be considered for people of Asian family origin who have recent-onset type 2 diabetes at a lower BMI than other populations).

Treatment pathway

BMI classification (kg/m^2)	Waist circumference*			Comorbidities present
	Low	High	Very high	
Overweight (25–29.9)	1	2	2	3
Obesity I (30–34.9)	2	2	2	3
Obesity II (35–39.9)	3	3	3	4
Obesity III (40 or more)	4	4	4	4
Treatment options				
1	General advice on health weight and lifestyle			
2	Diet and physical activity			
3	Diet and physical activity; consider drugs			
4	Diet and physical activity; consider drugs; consider surgery			
Notes:*for men, waist circumference of less than 94cm is low, 94–102cm is high and more than 102cm is very high. For women, waist circumference of less than 80cm is low, 80–88cm is high and more than 88cm is very high.				

6

Source: 'Obesity: the prevention, identification, assessment and management of obesity in adults and children' (NICE clinical guideline 189).

Patient's perspective

- Living with obesity can be a struggle
- Difficult to participate in certain activities – feel excluded
- Feeling of being judged by others – stigma
- Difficulties of losing weight and then putting it back on – vicious cycle
- Current support varies between regions – some areas offer lots through Tier 3 (covers specialist weight management) and 4 services (covers bariatric surgery)
- There needs to be a focus on the underlying mental cause of weight gain through psychological support
- The technology has a place in the current pathway – no other treatments address appetite or satiety

7

Source: Patient organisation submission – Helping Overcome Obesity Problems (HOOP)

Clinical expert view

- A lot of geographical variation in the care provided for overweight and obesity
  - There is a need to provide a more comprehensive and equitable service
- Current treatment options
  - Dietary advice and physical activity alone only results in a third of people achieving a sustained 5% weight loss
  - Bariatric surgery is highly effective but access is limited
  - Limited pharmacological options – orlistat use often limited by gastrointestinal adverse effects
- Naltrexone-bupropion provides a new option
  - Trials show reasonable efficacy in people without type-2 diabetes – ½ achieve a weight loss of 5% and ¼ achieve a weight loss of 10%; weight loss is slightly lower is people with type-2 diabetes
  - Overall cardiovascular risk factors (blood glucose / HbA1c and lipids) improved more with active treatment than placebo in the trials, but blood pressure and pulse did increase slightly – reassurance from the LIGHT study (assesses cardiovascular outcomes with naltrexone-bupropion) suggests no increase in cardiovascular events events
- Technology could be used in tier 3 specialists clinics and primary care
  - No significant extra burden to the NHS to provide technology

8

Source: see clinical expert submission statement from University of Liverpool and Aintree University Hospital NHS Foundation Trust

Decision problem

	NICE scope	Company submission
Population	Adults who have a BMI of: • ≥30kg/m^2 (obese) • ≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities	As per scope
Intervention	Naltrexone-bupropion prolonged-release (NB32)	As per scope
Comparator(s)	• Standard management without NB32 • Orlistat (prescription dose)	As per scope
Outcomes	• BMI • Weight loss • Percentage body fat • Waist circumference • Incidence of Type 2 diabetes • Cardiovascular events • Mortality • Adverse effects of treatment • Health-related quality of life	BMI missing Company considered BMI within the economic modelling, but it was not explicitly provided as a clinical outcome of the 4 COR trials as this was not a pre-defined endpoint
Subgroups	People with Type 2 diabetes	As per scope

Source: section 1.1 of the company's submission.

ERG comments

Comparators – it is not clear what is meant by 'standard management without NB32'

Outcomes BMI and percentage body fat are not reported in the company's submission. The data on cardiovascular events are also limited.

Pivotal randomised placebo-controlled trials

Trial name	Population	Intervention	Co-Primary Outcomes
COR-I **Phase III multicentre, double-blind** **Location: USA**	Adults with uncomplicated obesity or who were overweight with dyslipidaemia or hypertension	• Naltrexone 32mg per day + bupropion 360mg per day (NB32) • Naltrexone 16mg per day + bupropion 360mg per day	Mean percent change in body weight and proportion of patients with ≥5% decrease in body weight at week 56
COR-II **Phase III, multicentre, parallel-arm, double-blind** **Location: USA**	As above	NB32	Mean percent change in body weight and proportion of patients with ≥5% decrease in body weight at week 28
COR-BMOD **Phase III multicentre, double-blind** **Location: USA**	As above	NB32 + intensive behaviour modification (BMOD)	Mean percent change in body weight and proportion of patients with ≥5% decrease in body weight at week 56
COR-DM **Phase III multicentre, double-blind** **Location: USA**	Adults with T2DM and BMI ≥27 and ≤45kg/m^2	NB32	As above

Note: NB32 and placebo are all given as adjunct to standard management (SM) or intensive SM [BMOD] in COR-BMOD. COR, Contrave obesity research; DM, diabetes mellitus; BMOD, intensive behaviour modification; T2DM, Type 2 diabetes mellitus

10

Source: section 4.2 of company's submission for full details.

Other RCTs

The company submission also presented limited results for two other trials:

- IGNITE – NB32 + comprehensive life style intervention vs. usual care to study the percent change in weight from baseline. This study had a low number of participants.
- NB-CVOT – included participants with increased CV risk factors to study if NB32 reduced the time-to a major adverse cardiac event. This study was terminated.

To note: No direct head-to-head designed RCTs trials were presented by the company.

Standard management definitions in the COR trials

True to practice in England?

- **COR I and II** '*Participants were encouraged to increase physical activity, with a prescription for walking starting with at least 10 minutes on most days of the week, and increasing this gradually to 30 minutes on most days of the week throughout the study. They were encouraged to lose weight and maintain weight loss, and were encouraged to follow the prescribed programme (as described). Participation in any other weight loss programme was not permitted. The use of meal replacements (such as Slim Fast or Weight Watchers) was discouraged, but occasional use did not necessitate withdrawal from the study. The prescribed exercise could be performed in a gymnasium or health club.*'
- **COR-DM** *same as COR I and II but 'participants were encouraged to walk at least 30 mins in the first instance'*
- **COR-BMOD** '*consisted of group meetings lasting 90 minutes weekly for the first 16 weeks, every other week for the next 12 weeks and monthly thereafter. They included instructions to consume a balanced deficit diet and to increase to 180 min/week of planned, moderately vigorous, physical activity*'

11

Source: see company's clarification response

ERG comments on the trials

- Four main COR trials are of high quality but no trials directly compared NB32 with orlistat
- All trials conducted in the USA
  - Standard care may be different to that in England – regimen seen in COR-BMOD may be more reflective to that seen in England (group meetings mimics weight loss programmes)
  - Majority of participants were female – in England males are more likely to be overweight or obese, 68% vs 58%, respectively in 2015
- Overweight (approximately 2%) and Asian people are not well represented in the trials
- COR-I, -BMOD and -DM measure the primary outcomes at 56 weeks but there is no information on maintenance of weight loss after this time
- Prior use of orlistat was an exclusion criterion in all 4 COR trials so the effect of NB32 after orlistat has failed has not been examined

12

Source: see section 4.2 of the ERG report

Clinical trial results - mean % change in body weight from baseline

Modified intention-to-treat population (mITT) using last observation carried forward (LOCF)

Trial name		Baseline mean kg (SD) / Assessment point mean kg (SD)	Difference in Least Square (LS) Mean (95% CI), p-value NB32 vs Placebo
COR-I	NB32 (n=471)	100.2(16.3)	-4.8 (-5.6, -4.0) <0.001
		94.2(17.4)	
	Placebo (n=511)	99.3 (14.3)	
		98.0 (15.2)	
COR-II	NB32 (n=825)	100.7 (16.7)	-4.6 (-5.2, -3.9) <0.001
		-94.2 (17.6)	
	Placebo (n=456)	99.3 (16.0)	
		97.2 (16.2)	
COR-BMOD	NB32 +BMOD (n=482)	100.7 (15.4)	-4.2 (-5.6, -2.9) <0.001
		91.0 (17.1)	
	Placebo + BMOD (n=193)	101.9 (15.0)	
		96.4 (17.1)	
COR-DM	NB32 (n=265)	106.4 (19.1)	-3.3 (-4.3, -2.2) <0.001
		101.0 (19.7)	
	Placebo (n=159)	105.0 (17.1)	
		103.0 (17.3)	

13

Source: see section 4.7 of the company's submission for full results

Modified Intention-To-Treat population

The results are presented for a modified intention-to-treat population. The participants who were included in the modified ITT analysis had to meet the following 3 criteria:

- A baseline body weight was recorded
- Patient was randomised
- A post-baseline body weight was recorded while patient was on treatment

The CHMP had concerns with mITT population and the last observation carried forward (LOCF) method due to high drop out rates. A sensitivity analysis was performed with the ITT population (at least one post-baseline weight measurement) using the baseline observation carried forward (BOCF) method and WRIM (weight regain imputation method – assumes a regain of 0.3kg per month following study withdrawal) to address the concerns and found the sensitivity analyses substantiated the results of the primary efficacy analysis.

mITT vs ITT (Last observation carried forward [LOCF] vs baseline observation carried forward [BOCF])- source: see company's clarification response

At clarification the ERG asked the company to provide a re-run of the analysis with the ITT population and other imputation methods. The results show a small increase in the odds ratios with BOCF and WRIM compared with the base case mITT-LOCF method and a decrease in the mean differences.

Below presents the company's rationale for the mITT over other methods:

'The mITT population was defined in the company's submission as patients who had at least one post-baseline weight measurement obtained while the patient was still taking study medication, with missing data imputed using the last observation carried forward (LOCF) method. To derive weight loss outcomes for patients beyond 16 weeks, it was required to establish the cohort of patients that responded at 16 weeks. Regardless of population utilised, the subset of patients that responded to treatment at Week 16 is the same. As such, the only weight loss outcomes required for the model that could utilise the ITT populations are those at the Week 16 assessment.

However, within the economic analysis, weight loss outcomes were separated by those who respond to treatment and those who do not, with a randomly sampled number utilised to determine whether the patient is a responder or a non-responder. By utilising weight loss outcomes for patients with no further observations from baseline (as is implied by considering the ITT populations over the mITT population), the proportion of primary assessment non-responders will be over-estimated as the analysis will automatically assign all patients with no further measurements as non-responsive.

The use of BOCF to impute missing data may result in further overestimation of the number of non-responders, as a patient that discontinues from the study post baseline is also assumed to have had no change in weight; a patient that discontinues towards the end of the study would therefore be assumed to have received no treatment effect, which is clearly unlikely.

Although data imputation using LOCF avoids this issue, it is acknowledged that patients are likely to regain weight post discontinuation of treatment (that is, standard management and adjunctive therapy), but this has been considered in the model by applying a linear regain period of 3 years.'

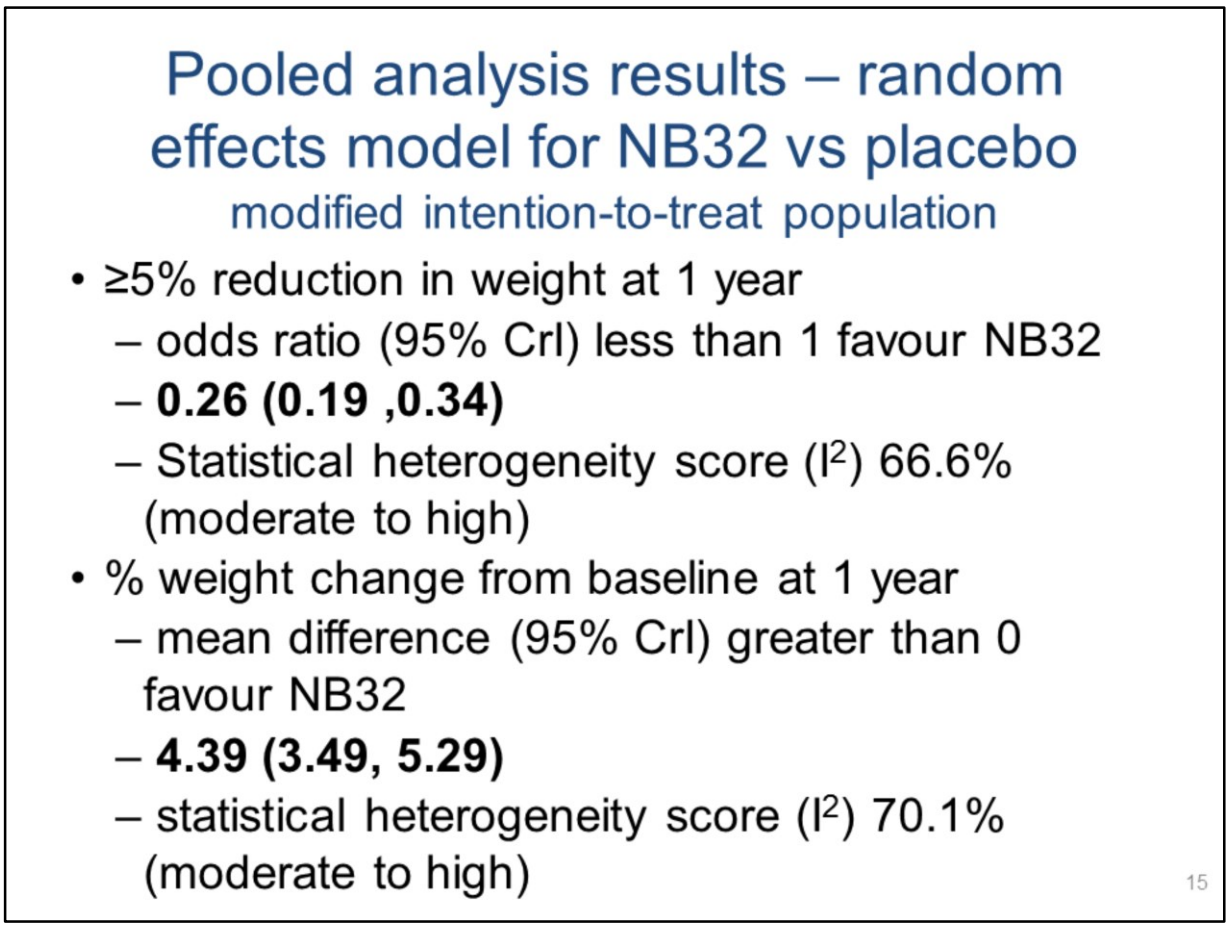
Clinical trial results - ≥5% decrease in bodyweight from baseline

Modified intention-to-treat population (mITT) using last observation carried forward (LOCF)

Trial name		N (%), 95% CI	odds ratio (OR) (95% CI) , p-value (Higher odds favour NB32)
COR-I	NB32 (n=471)	226 (48.0%), 43.5, 52.5	4.9 (3.6, 6.6), <0.001
	Placebo (n=511)	84 (16.4%), 13.2, 19.7	
COR-II	NB32 (n=825) to week 28	459 (55.6%), 52.3, 59.0	6.6 (5.0, 8.8), <0.001
	Placebo (n=456)	80 (17.5%), 14.1, 21.0	
COR-BMOD	NB32 +BMOD (n=482)	320 (66.4%), 62.2, 70.6	2.9 (2.0, 4.1), <0.001
	Placebo + BMOD (n=193)	82 (42.5%), 35.5, 49.5	
COR-DM	NB32 (n=265)	118 (44.5%), 38.5, 50.5	3.4 (2.2, 5.5), <0.001
	Placebo (n=159)	30 (18.9%), 12.8, 25.0	

14

Source: see section 4.7 of the company's submission for full results.



Source: see section 4.9 of the CS for full pooled analysis

Note: pooled meta-analysis methodology

- Included the four pivotal trials – all had similar patient populations
- COR-I, II and BMOD excluded people with T2DM and BMOD had a more intensive standard management regimen
- Frequentist pairwise meta-analysis to assess
  - At least a 5% reduction in weight at 1 year from baseline (the 1-year time point ranged from 52 to 57 weeks). This was a dichotomous outcome. Measured as odds ratio (ORs)
  - Mean % weight change from baseline at 1 year (the 1-year time point ranged from 52 to 57 weeks)
- Random-effects model chosen as it allows to capture between-trial heterogeneity). This was a continuous outcome. Measured as a mean difference (MDs

Results interpretation from the company's submission

≥5% reduction in weight at 1 year

The pooled results shows the chance of an event (odds of an equal or greater than 5% reduction in weight) is 74% more likely with NB32 than placebo. To note; this was expected across all trials as the patient populations and treatments in the trials were similar (as stated in

the company's submission). In the COR-DM and COR-BMOD trials, differences between NB32 and placebo were less pronounced (compared to COR-I and COR-II). Nevertheless, results were still significantly in favour of NB32. For COR-BMOD in particular, the higher OR reflects that more people on placebo lost at least 5% of their initial weight due to the more intensive behaviour modification program relative to people on placebo in the other studies.

% weight change from baseline at 1 year

The pooled result is presented as 'weighted' (because they are pooled – but no weights given to any of the trials) mean differences which show people who received placebo had a significantly smaller % reduction in weight (at 1 year compared to baseline) versus NB32, for all COR trials. To note, 'The COR-DM trial produced lower mean differences of response compared to the COR-I and COR-II trials for placebo versus NB32. The mean difference in the COR-BMOD trial is also lower than the COR-I and COR-II trials for placebo versus NB32; however, there is more uncertainty around this estimate. The I^2 value indicates moderate-high heterogeneity, which is likely to be due to the lower MD observed in the COR-DM trials compared to the COR-I and COR-II trials'.

Adverse Events in the COR trials

- In all 4 trials there were more treatment-emergent adverse events in the NB32 arm compared to placebo
  - Ranging from 57.1% to 76.5% in the COR trials for NB32
- Common AEs across the trials were GI (nausea and constipation) and CNS related (headache and dizziness) – nausea was the most common AE leading to discontinuation, NB32 vs PBO:
  - 19.5% vs 9.8% in COR-I
  - 24.3% vs 13.8% in COR-II
  - 25.7% vs 12.5% in COR-BMOD
  - 29.4% vs 15.4% in COR-DM
- Cardiovascular effects (naltrexone) and psychiatric effects (bupropion) are two AEs of concern outlined in the SmPC
  - No significant numbers of cardiovascular (e.g. increased blood pressure) or psychiatric (e.g. suicidal thoughts) effects reported across the 4 pivotal trials

16

Treatment-emergent adverse events - defined as events that first occurred or worsened during double-blind treatment (i.e. a new event or an exacerbation of a pre-existing condition) with an onset date after study drug administration and within 7 days of the last confirmed dose date

ERG comments

ERG note that there are a greater proportion of gastrointestinal events, particularly nausea, in the NB32 groups across the trials and there were a large number of withdrawals from the treatment groups compared to placebo (usually due to GI side-effects).

ERG comments on the COR trial results

- Using a modified-Intention-to-treat (mITT) population could lead to bias:
  - mITT included people who had at least one post-baseline measurement (approximately 20% of patients excluded). Reasons for discontinuation could relate to efficacy or safety of NB32
  - True ITT should be used – results for NB32 vs placebo are still significant but more modest
- Inappropriate to pool COR trials because of clinical & statistical heterogeneity:
  - Results from the separate analyses for patients with and without diabetes are preferred
  - COR-BMOD not suitable to be pooled with the other COR trials as standard management was more intensive and greater weight loss was achieved. Placebo arm in COR-BMOD had results approaching the intervention arm of the other trials
  - Use of COR-II to derive treatment effect beyond 28 weeks is inappropriate because NB32 participants with ≥5% weight loss at visits between weeks 28 and 44 were re-randomised
- Large drop out rates due to adverse events with NB32 in the trials (up to 50%)

17

Source: see section 4.2.4 of the ERG report

Indirect treatment comparison (ITC) – NB32 vs orlistat

- Company presented an ITC with placebo as the common comparator (using pooled results from meta-analyses for NB32 vs placebo)
- A Bayesian network meta-analysis (NMA) was performed to assess:
  - ≥5% reduction in weight from baseline at 1 year
  - mean % weight change from baseline at 1 year
- Analyses were presented for:
  - People with T2DM
  - People without T2DM
  - All trials regardless of T2DM
- Random effects model used only for all trials, regardless of T2DM. Fixed models used for T2DM and no T2DM subgroups
  - Sensitivity analysis performed to explore heterogeneity (differences in intensity of BMOD and lead-in periods) in trials and found consistent results to the base case

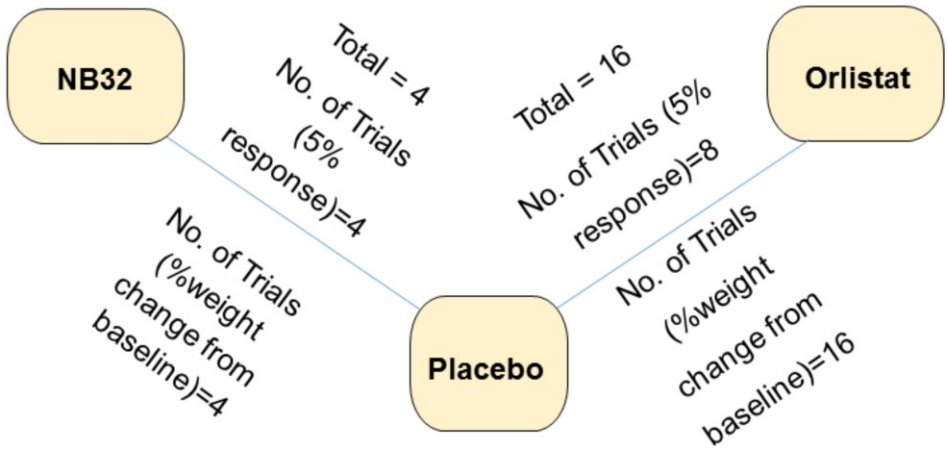
18

Note:

Company only presented Bayesian NMA results in the main submission..

Random effect results are not presented for the T2DM and non-T2DM analyses, as the models failed to update effectively in WinBUGS using the recommended priors, likely due to the low number of studies.

ITC – Network of evidence



19

ITC base case results – NB32 vs orlistat
Modified ITT with LOCF

Trials	≥5% reduction in weight (1 year), odds ratio (95% CrI)	Mean % weight change (1 year), mean difference (95% CrI)
Trials with people with T2DM (Fixed effects)	1.09 (0.63 to 1.88)	0.21 (-0.87 to 1.30)
Trials excluding people with T2DM (Fixed effects)	0.77 (0.61 to 0.96)	1.13 (0.44 to 1.80)
All trials regardless of T2DM (Random effects)	0.80 (0.51 to 1.28)	1.39 (-0.08 to 2.82)
Note, Odds Ratio less than 1 and Mean Differences greater than 0 favour NB32. *, Company presented results of the pooled meta-analysis as log odds ratios. These are converted back to natural scale for comparative purposes		

20

Source: see section 4.10 of company's submission for full results and sensitivity analyses and section 4.4 of the ERG report.

Results interpretation

For people with T2DM – the results show that there is may be no difference between orlistat and NB32 (OR; 1.09 [0.63, 1.88] for equal or greater than 5% reduction in weight and MD; 0.21 [-0.87, 1.30] for mean % weight change from baseline). There is a greater effect in favour for NB32 in the trials excluding people with T2DM and a 5% chance there may be no difference in effect in all trials.

Note: Sensitivity analyses were performed to assess risk of bias for:

SA1: Trials with 'high' comorbidities were excluded – for all trials regardless of T2DM – to assess the heterogeneity in participants via effects of weight loss in trials where a large proportion of patients had comorbidities, a sensitivity analysis was performed excluding trials where ≥75% of patients had at least one comorbidity (hypertension, dyslipidaemia, or T2DM)

SA2: Trials with lead-in periods were excluded – for all 3 analyses – 11 trials included a lead-in period (a period where patients receive non-active therapy a to assess tolerability) compared to the 4 COR trials.

SA3: Trials with intensive BMOD were excluded – for trials excluding T2DM and all trials regardless T2DM – Because standard management was considered an additive benefit, the effect of intensive behaviour modification received in COR-BMOD was investigated compared to the other trials where only standard management was received

- Results increased in favour for NB32 for a mean % weight change and for a ≥5% reduction in weight, at 1 year, for the analysis for no T2DM trials

SA4: Trials with lead-in periods or intensive BMOD were excluded – for all trials regardless of T2DM only

ERG comments on ITC

- ERG considers Bayesian NMA methodology is appropriate:
  - Agrees that only fixed models are presentable for T2DM and no T2DM subgroups and are likely to be more reliable
  - Appropriate sensitivity analysis was explored by the company
- Full comparisons not considered by the company:
  - NB32 plus standard management (SM) vs intensive SM
  - NB32 plus intensive SM vs orlistat plus intensive SM
- Additional work by ERG:
  - Using mITT data is main concern – mITT population in NB32 trials (21.9% of patients excluded) very different from in orlistat trials (1.6% excluded)
  - Two additional analyses provided by the ERG
    1. Results based on ITT populations for the NB32 trials
    2. Comparison of studies with intensive BMOD

Results used in ERGs preferred analysis

21

Source: see section 4.5 of the ERG report for full results for the two additional analyses.

ERG preferred ITC analyses – using Bucher method for ITC and ITT-baseline observation carried forward analysis (ITT-BOCF) and no pooling of NB32 trials - NB32 vs orlistat

Trials	≥5% reduction in weight (1 year), OR (95%CrI)	for mean % weight change (1 year), MD (95% CrI)
People with T2DM	1.59 (0.89 to 2.79)	-1.21 (-2.30 to -0.11)
People without T2DM	0.61 (0.31 to 1.22)	1.11 (-0.39 to 2.63)
Intensive behaviour modification	1.86 (1.30 to 2.66)	-2.09 (-3.53 to -0.65)
Note, Odds Ratio less than 1 and Mean Differences greater than 0 favour NB32		

22

ERG preferred analysis

The ERG prefer not to pool the NB32 trials and therefore have not presented any results for 'all trials regardless of T2DM' subgroup as this would mix people with T2DM from COR-DM and non-T2DM from the other trials.

Trials that use intensive BMOD are also excluded and considered separately, which means the results for T2DM only results will be the same as the company's ITT results but the results for where T2DM is excluded will change.

Using the ITT population shows that the positive effect for NB32, in the T2DM only subgroup, has all but disappeared compared to the mITT results in the company's submission.

When a comparison is made between the COR-BMOD and XENDOS (orlistat as an adjunctive treatment to intensive BMOD) trial, in the intensive BMOD subgroup, the results show a superior positive effect for orlistat compared to NB32.

Bucher method

Common method to perform an indirect comparisons using a common comparator to estimate the point estimate and its confidence. Helpful in analysing subgroups.

Key issues: clinical effectiveness

1. **Positioning**
   - What is the expected positioning of naltrexone-bupropion in the treatment pathway: alternative to orlistat vs 2nd line to orlistat? What are the relevant comparators? Are different types of behaviour modification, such as more intensive forms, relevant?
2. **Population**
   - should effectiveness be considered in a mixed population (overweight and obese) with and without Type 2 Diabetes Mellitus (T2DM)?
3. **Effectiveness of naltrexone-bupropion (NB32) vs placebo**
   - What is the appropriate analysis: Intention to treat (ITT) or modified ITT (mITT)? Implications of large drop out rate, and how to deal with this analytically: Last observation carried forward, Baseline observation carried forward, Weight regain imputation? Should the COR trials be pooled?
4. **Effectiveness of NB32 vs orlistat**
   - Which trials should be used in the indirect treatment comparison of NB32 vs orlistat?
5. **Generalisability to NHS**
   - Is standard management in the COR trials generalisable? Is the patient population in the trials generalisable?

23

Cost effectiveness evidence

Key issues: cost effectiveness (1)

1. Model structure

- Model structure does not consider retreatment, behaviour modification, and bariatric surgery. What is the committee's view on the model structure?

2. Model implementation

- The ERG highlighted that the model is very slow to run, large variation in ICERs when different random numbers are used and small number of PSA runs, and BMI updated only when events occur. What is the committee's view on the validity of the model and robustness of the results?

3. Population

- Should the cost-effectiveness be considered in the entire population or in subgroups with/without T2DM? What are the characteristics of the population that should inform the model?

25

Key issues: cost effectiveness (2)

4. Modelling treatment

- What clinical data is appropriate to inform the model? Duration of effect: how fast is weight regained after treatment discontinuation? Treatment duration: is time on treatment appropriately modelled?

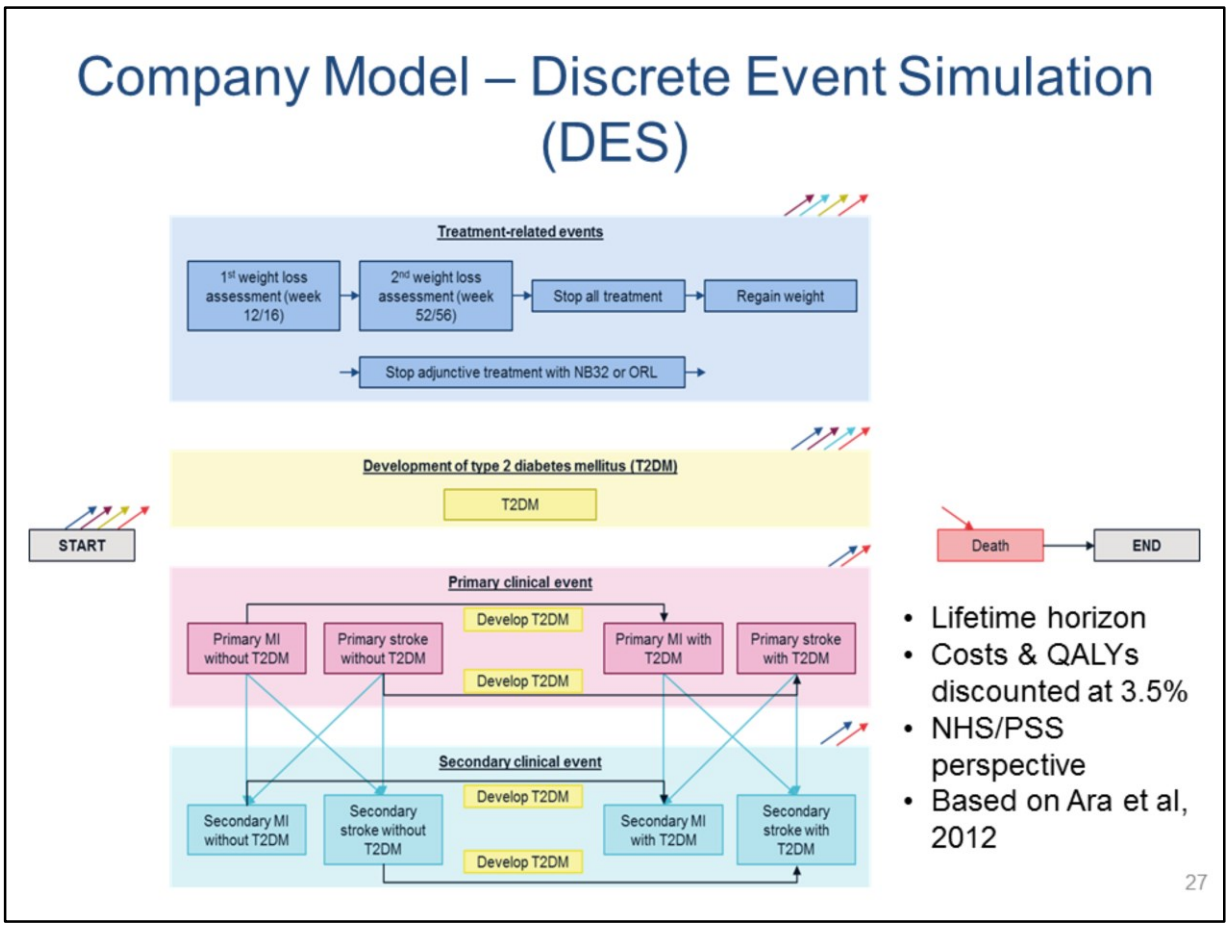
5. Utilities

- Is the Public Health England weight management tool appropriate to derive the utilities?

6. Innovation.

- Does the committee consider NB32 to be an innovative therapy?

26



Source: see section 5.2.2 of the CS for full model details

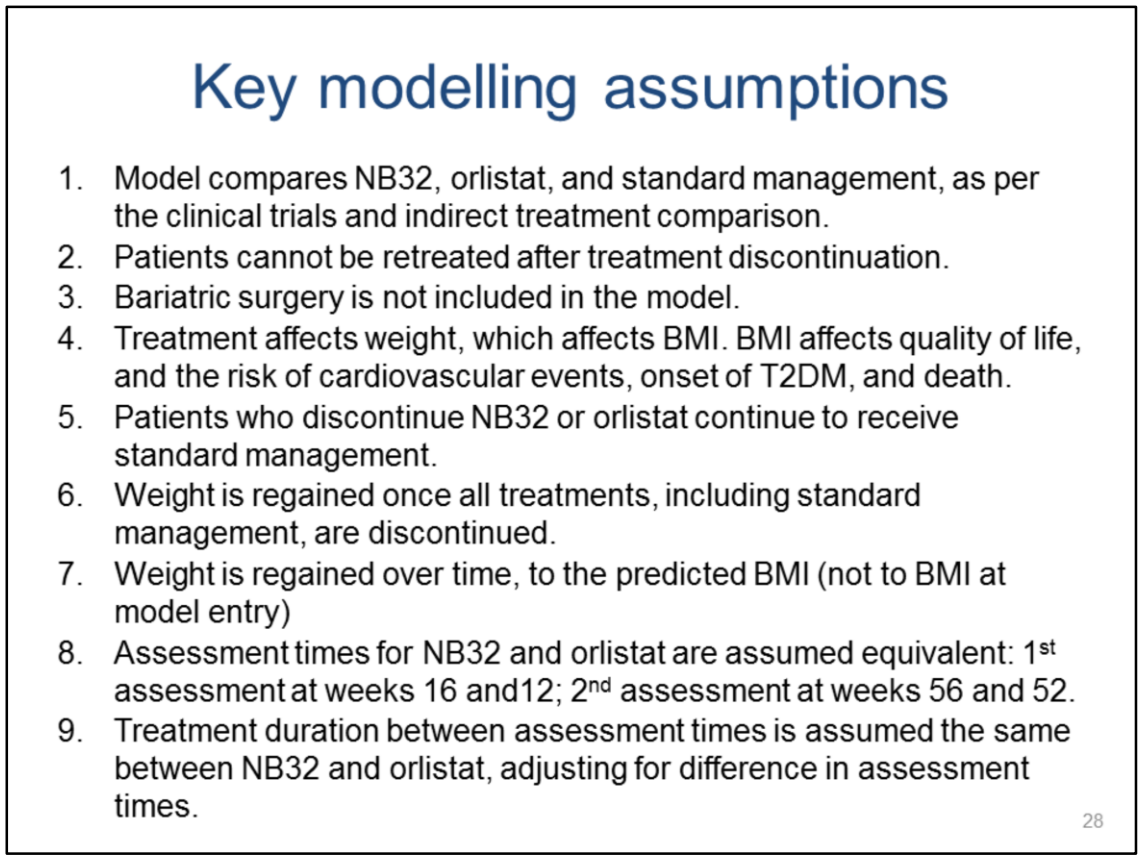
Company's economic model

- De novo analysis using DICE methodology – based on Ara et al (2012) model
- Individual patient simulation model (DES) – one patient (assigned baseline characteristics*) followed through to death three times before the next patient enters
- First patient run – assigned NB32 as adjunctive treatment (alongside standard management)
- Second patient run – assigned orlistat as adjunctive treatment
- Third patient run – assigned standard management only
- The patient progresses through the model and may experience various treatment or disease events which has consequences for patient utility and/or on health and social care costs
- Patient followed until death and lifetime costs and QALYs calculated

DICE methodology Source: Evidera website

- Disease process and its management defined as two fundamental aspects
  - 'conditions' that can exist and 'events' that can happen
- Each event has a set of consequences that are processed when the event occurs

- The level of each condition can change over time and is updated when an event occurs
- Conditions
    - Persist over time, have levels which affect event and conditions, many can be present at once
    - We are interested in time spent at a given level
- Events
    - Happen at a point in time, can affect other events and conditions' levels, many can happen at any time
    - We are interested in the number of events that happen



Source: section 5.2 of the company's submission, complemented with ERG report.

Company's Rationale for assumptions

Treatment discontinuation – clinical expert view

Weight regain - This assumption was made in the model built by Ara et al. For people who discontinue adjunctive therapy but continue to receive non-pharmacological standard management, weight regain was assumed to only commence when standard management was discontinued. Clinical expert opinion was sought to validate this assumption

Obesity-related clinical events - This assumption was made in the model built by Ara et al. It is expected that the incremental clinical impact of further cardiovascular events would be negligible, as the proportion of people who would experience more than two cardiovascular events in clinical practice is small

Baseline characteristics

Parameter	Mean value	Justification
Age	47.0 years	COR trial programme patient-level data
Female	79.0%	
Height	Female: 1.64 m Male: 1.78 m	
BMI	Predicted by natural history model; average of 33Kg/m^2	BMI trajectory model by Ara et al.
T2DM at baseline	33.2%	Ara et al.
Insulin use for T2DM patients	33.3%	Clinical opinion
Smoking status	Current: 7.0% Previous: 54.0% Never: 39.0%	Dare et al.
Statin use	79.3%	NB-CVOT study
History of angina	0%	Assumption – no data identified for overweight/obese patients
Other type DM	0%	

29

Source: see section 5.3.1 of the company's submission.

Sources for key data

Where possible, data were utilised from the COR trial programme, followed by the NB-CVOT study and then alternative data sources. Age, gender and height values were all derived using patient-level data from the COR trial programme.

BMI - For consistency with later model projections, BMI was derived at baseline using the BMI trajectory model by Ara et al. Use of this model ensures estimated changes in BMI over time are logical, given that following all treatment cessation, people are assumed in the base case analysis to regain weight linearly over a 3-year period until their projected BMI at this time.

T2DM - The proportion of people with T2DM at baseline was taken from Ara et al. as the majority of studies in NB32 and orlistat were either conducted in non-diabetics or only diabetics. Insulin use for diabetics was assumed to be 33.3%, in line with clinical expert opinion that diabetes treatment comprises of insulin for around a third of patients.

ERG comments on model and assumptions (1)

1. Modelling structure
   - Inability of model to incorporate re-treatment, behavioural modification treatment and bariatric surgery is a major limitation
2. Implementation
   - Model very slow to run. Simpler approaches (e.g. individual-level state transition model) may have been more appropriate.
3. BMI over time.
   - Model does not update BMI frequently enough (after year 1, on average updated once every 10.6 years).
4. Reasonable to use the Ara et al model as a starting point but issues on deviations

30

Source: see section 5.2 of the ERG report

ERG comments on model and assumptions (2)

5. Assumption on weight regain. Weight regain is a key assumption and driver in Ara model. Company deviated from assumption that patients would have regained weight to obtain their *baseline* BMI in 3 years and assumed instead that patients would have regained weight to obtain the *predicted* BMI in 3 years
   - ERG not satisfied with this deviation and prefer the assumption used in Ara
   - In response to clarification the company provided an analysis where BMI returned to baseline (ICER vs orlistat increased by £1,536)
   - Linear weight regain over 3 years implemented incorrectly (instantaneously at end)

6. Comparability of assessment times. Company model assumes weight loss with orlistat at weeks 12 and 52 is comparable to weight loss with NB32 at weeks 16 and 56 but no justification given.

31

ERG comments on baseline characteristics

- Agree with using data from COR trials as effectiveness estimates are derived from this population but other baseline characteristics are questionable
- Baseline BMI is vastly underestimated in the model compared with the trials
  - Therefore utility, and time to T2DM, CV events and death could be overestimated as BMI is included as a predictive factor
- Other baseline characteristics underestimated in the model vs trial data include:
  - % current smokers (7% vs 9-11% in trials)
  - % receiving anti-hypertensive medication (0% vs 15-63% in trials)
  - ERG disagrees with assumption that no patients had a history of angina and/or diabetes other than T2DM – model results therefore not representative
- Some baseline characteristics overestimated:
  - % with T2DM (33.2%). Health Survey for England data suggests 14-15% (overweight and obese)
  - % on statins (79% vs 8-13% in trials)

32

Source: see section 5.2 of the ERG report and table 5.4 of the ERG report.

ERG preferred values

Age 47 (model), T2DM; 53.8 non-T2DM; 44.7 (ERG) – the percentage of people with T2DM and obese are greater than that presented by the company – the model should reflect the population in the trial which is a poor representation of people who are overweight (percentages therefore reflect people who are obese)

Current smokers 5.7% (model), 1.6% (ERG)

Receiving anti-hypertensives 0% (model), T2DM; 47.9% non-T2DM; 15% (ERG)

Statin use 80.4%(model), T2DM; 47.6% and non-T2DM; 10.4% (ERG)

Receiving aspirin 0% (model), 10.9% (ERG)

Clinical data used in the model

- **NB32 and SM**:
  - Proportion of responders at weeks 16 and 56, and change in body weight from pooled COR trials (modified-ITT analysis with LOCF).
- **Orlistat**:
  - Proportion of responders at weeks 12 and 52, and change in body weight from indirect treatment comparison (modified ITT population).
- **Time to treatment discontinuation (TTD) for NB32/orlistat**
  - 3 periods: up to week 12/16, between week 12/16 and week 52/56, and from week 52/56 onwards.
  - Periods up to week 52/56 based on pooled COR trials; period 56 onwards based on NB-CVOT trial.
  - Orlistat was assumed the same as NB32, with adjustments for different assessment times, due to lack of data for orlistat.
- **BMI over time.**
  - BMI over time predicted based on sex and age from the Ara et al model.
- **Impact of weight on events**
  - Changes in body weight , converted to BMI, were used to predict development of T2DM, CV event (stroke or MI) and death using parametric time-to-event models (Weibull) retrieved from Ara et al.

33

Source: see section 5.3 of the company's submission and section 5.2.6 of ERG report

To estimate the percentage weight loss for orlistat at the primary assessment point; the estimate was taken from the ITC but again the value at 1 year was assumed equivalent to week 16. At the secondary assessment point The weight loss for NB32 at Week 56 was assumed comparable to the weight loss for orlistat at Week 52, given the lack of a 4-week titration period for people treated with orlistat.

To estimate weight loss for people on standard management at first response at Weeks 12 and 16 were derived using available COR trial programme patient-level data. This data were not separated by response as standard management would not be assessed for response. At second response assessment weight loss at Week 52 and 56 were also derived using the ITC, as with orlistat. However, as people treated with standard management alone are not subject to the same response-based treatment stopping rules as those treated with NB32 or orlistat, the base estimate from which the ITC was applied was taken to be the estimated weight loss for NB32 at Week 56 regardless of response at Week 16

Time –to -Treatment discontinuation (TTD) methodology

- Treatment initiation to first assessment point
  - KM estimates produced from the COR trials to estimate number of people still on

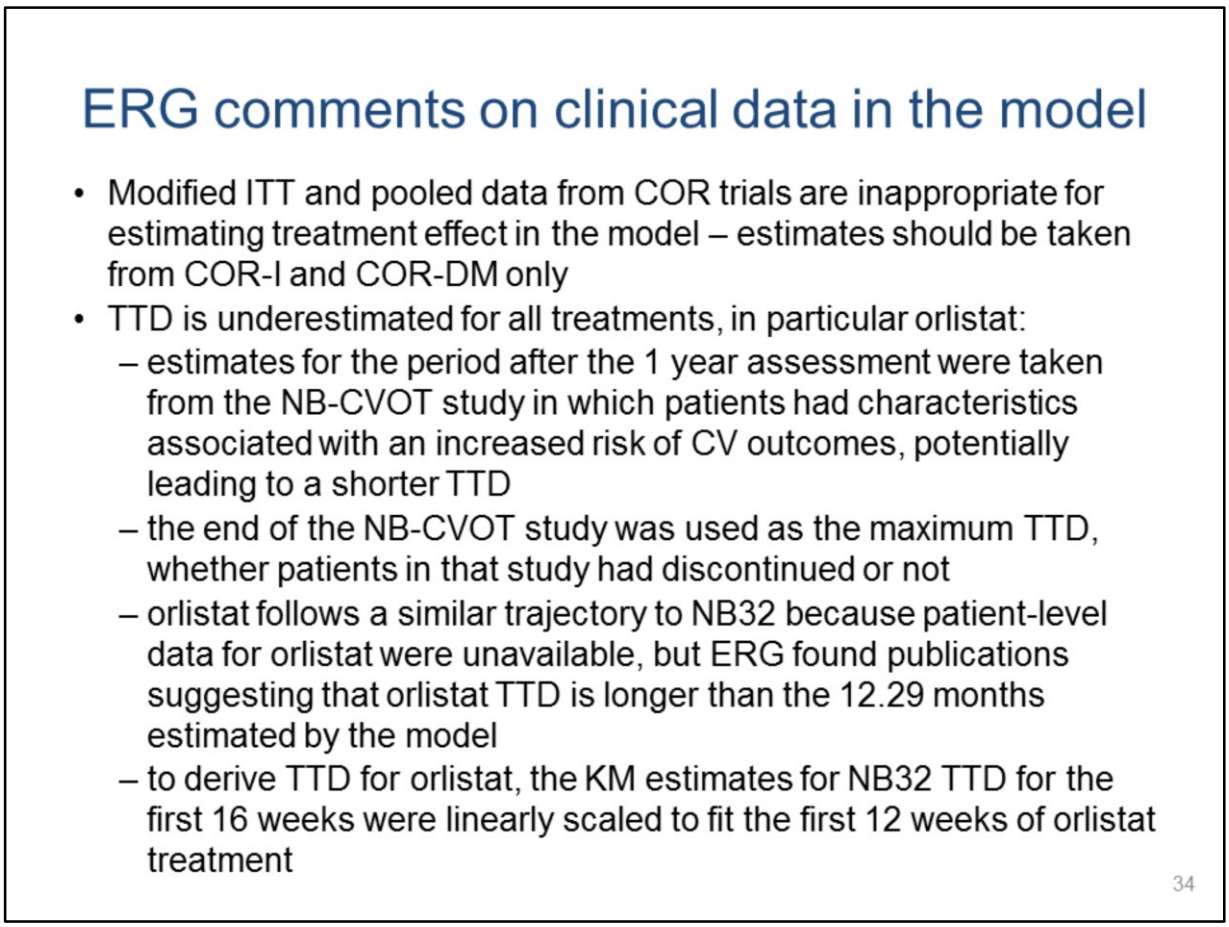
adjunctive treatment – 67.2% continued until week 16 on NB32

- No comparable data available to estimate orlistat so NB32 (week 16) KM data linearly scaled to estimate orlistat (week 12)

- Treatment initiation to second assessment point
  - Analysis as per first assessment point but included people who achieved a weight loss of at least 5% (vs baseline)
  - Estimates generated for people on treatment to week 56 for NB32 and week 52 for orlistat
  - 86.1% of responders from week 16 continued until week 56 on NB32
- Beyond second assessment point
  - NB-CVOT trial used to inform long-term on treatment duration to week 156 for NB32 and week 152 for orlistat
  - For SM the KM estimated curves were used up to the second assessment point (as part of combined treatment) and then extrapolated by tagging on the KM curve for SM from NB-CVOT (as stand alone) to inform long-term duration of treatment

NB-COVT trial

A *'a Phase IIIb, multicentre, randomised, double-blind, placebo-controlled trial to assess the occurrence of MACE (major adverse cardiac events) in overweight or obese patient'.* 'The main differences between the NB-CVOT trial and the COR trials is that participants were all at increased risk of adverse cardiovascular outcomes. Furthermore, the trial incorporated a lead-in period. During the lead in period 1,490 patients discontinued. Of these 543 discontinuations were due to adverse events'. (ERG report page 57)



Source: see section 5.2 of the ERG report.

ERG comments on clinical effectiveness estimates

Time-to-treatment discontinuation (TTD)

Inappropriate to derive TTD from the four pooled COR trials – ERG believe TTD may be different for those receiving intensive BMOD and for people with or without T2DM

TTD should be modelled separately for T2DM and non-T2DM subgroups

Company's model run revealed a mean TTD of 13.32 months, 12.29 months and 17.16 months for NB32, orlistat and for SM respectively - ERG thinks that these may be under-estimates

Proportion of responders with weight loss >5%

Discrepancy found between mean OR in the model compared to that in the company's submission (1.13 vs 1.09 – both in favour of orlistat)

Inappropriate to use mITT and pooled COR trials for the proportion of responders (ITC estimates based on pooled COR results)

Mean change in body weight

Again inappropriate to use a mITT and pooled COR results to derive the mean change in body weight

Assuming on treatment effect of weeks 12 and 52 with orlistat is appropriate to compare to week 16 and 56 with NB32 – company justification (first 4 weeks of treatment with NB32 is a titration period) is inappropriate as people still lose weight in this period

Weight loss at 12 weeks with orlistat is derived using MD from the ITC

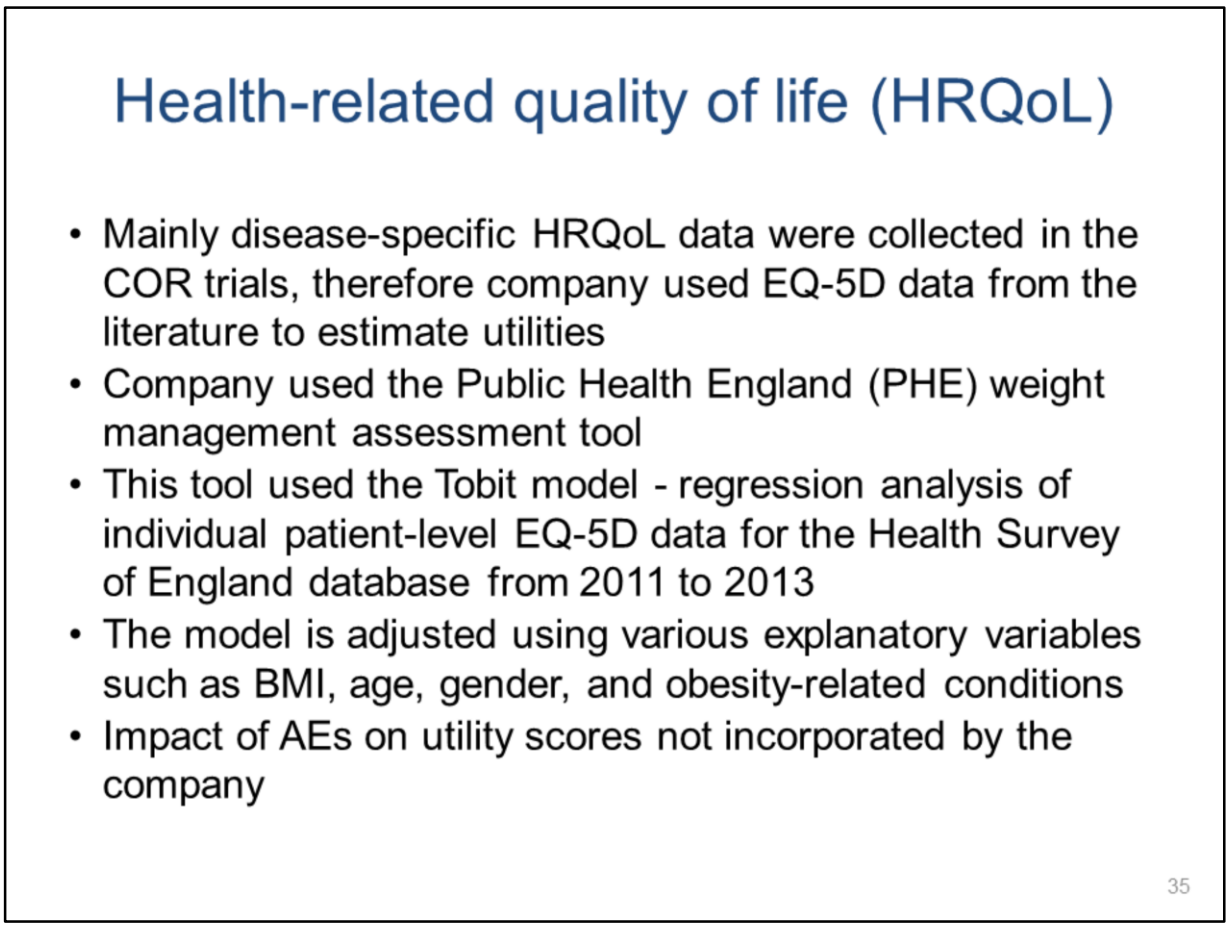
This is an absolute measure which varies according to the magnitude of weight loss

Because absolute weight loss at primary assessment being smaller than at secondary assessment at 1 year, applying the absolute MD at 1 year for NB32 would underestimate the weight loss for people treated with orlistat

ERG adjusted MD to relative risks in their preferred base case

Risk of obesity related events and natural history BMI model

ERG consider it appropriate to use the model reported in Ara et al but are concerned with using a lower BMI (reported in Ara natural history BMI model) than that found in the COR trials



Source: see section 5.4 of the CS table 60 for the regression coefficients used in the model and company's clarification response question B11 for the utility values.

Rationale for not including adverse event utility decrements

Expert opinion indicated that people on NB32 have a HRQL benefit over people on orlistat as a result of AE differences. 'While the side effects associated with NB32 are similar to those associated with many common drugs, the lower digestive tract AEs associated with orlistat can be particularly unpleasant for patients. The company reported that details of orlistat AEs from the clinical trial literature and publicly available regulatory documents are not sufficient to make appropriate trial-data comparisons between NB32 and orlistat adjunct therapies. Also the HRQL implications of the orlistat and NB32 AEs for obese and overweight people are poorly understood, and in some cases overlap and interact with obesity comorbidities. Therefore the company took the simplifying assumptions of no difference which it highlights is a conservative one.

Tobit model vs Ordinary Least squares (OLS)

A Tobit model (often termed a "censored [regression] model") is specifically designed to accurately reflect the distribution of data where censoring is known to apply – for example, at a lower or upper bound (or both). EQ-5D-3L utilities using a UK tariff . The Tobit model aims to estimate the proportion of people that are located at each of these bounds, and utilises these

within the estimation of the overall utility. In short, the Tobit model acknowledges the censoring limits and treats utilities at these limits separately to those in between. An OLS model simply fits a standard regression model to observed EQ-5D data without considering the censoring limits – source clarification response to question B11b.

ERG comments on HRQoL

- PHE model does not appear to be published in a peer-reviewed journal
  - Limited validity information on the model
- Concern that the estimates have no face validity
  - In response to clarification the company compared the values to those of the general UK population and the ERG was satisfied these showed face validity
- ERG agree that the Tobit model is more appropriate than ordinary least squares (OLS)
  - OLS disregards upper and lower bounds commonly used for estimating utilities
- ERG questioned company's claim for not including AE utility decrements
  - In response to clarification the company provided an analysis with utility decrement of 0.05 for all AEs over 1 week
  - The ICERs versus orlistat and SM increased by £188 and £87, respectively
  - ERG was satisfied that the impact of AEs on HRQoL is likely to be small

36

Source: see section 5.2.8 of the ERG report.

Resource use – included costs

- Drug acquisition costs
  - NB32 £73.00 per pack (112 tablets), Orlistat £18.44 per pack (84 capsules)
- No administration costs
- NHS resource use associated with medical monitoring
  - GP visits, nurse visit and blood tests
- NHS resource use associated with co-morbidities
  - Adapted from Ara et al.
  - Costs inflated from 2009 to 2015
- NHS resource use associated with managing AEs
  - Calculated from COR-I trials assuming one GP visit for NB32 and SM – orlistat assumed equivalent to NB32
    - NB32, £1.69/week; orlistat, £1.69/week and SM, £0.81/week
  - Outpatient costs according to disease area
- Drug wastage associated with NB32 not considered in base-case model

37

Source: see section 5.5 of the company's submission for full details and breakdown of the costs included in the model.

ERG comments on resource use

- Unclear why a GP visit was included at week 52 for SM
  - ERG removed this cost in its base case
- Unclear why company assumed only a single GP visit for each AE
  - Assuming outpatient costs increases ICER vs orlistat by £4,408
- Questionable whether assuming the same AE costs for orlistat as calculated for NB32 is appropriate
  - No direct safety evidence comparing the drugs
- Unclear why only COR-I was used to derive rate of AEs
  - COR-DM could have been used to inform rates for people with T2DM
- Excluding drug wastage is not a conservative assumption
  - ICER compared to orlistat increased by £3,426 when it is included

38

Source: see section 5.2.9 of the ERG report and section 5.5 of the company's submission.

Company base case results (deterministic)*

Technology	Total		Incremental		ICER (QALYs)	
	Costs	QALYs	Costs	QALYs	Versus baseline (SM)	Incremental
SM	£6,519	15.36				
orlistat	£6,814	15.41	£294	0.05	£5,538	£5,538
NB32	£7,563	15.44	£750	0.03	**£13,647**	**£32,084**

*The probabilistic analysis shows a similar ICER for NB32 versus standard management (£13,936) and a higher ICER for NB32 versus orlistat (£36,405)
Note, results rounded to 2 decimal places

39

Source: see section 5.7 of the CS.

To produce results in the base case a sufficient number of people are required to be run such that the model results converge to a consistent value. To establish how many patient profiles are required to produce stable model results, a diagnostic exercise was carried out. For total QALYs convergence occurred around 500 runs but was much larger for total costs. It was deemed that around 1,000 sample runs was sufficient to product stable base case results.

The company stress here that the results hinge on many conservative assumptions applied to the model:

- obesity-related health conditions the analysis considers are limited to MI, stroke and T2DM
- The blindness of the analysis to many cost and health benefits of weight loss means that the cost-effectiveness of more effective alternatives is inherently underestimated
- Model estimates for orlistat adjunct therapy are further limited by the key assumptions required to estimate the relative effectiveness of orlistat versus NB32 or standard management alone, and treatment duration for orlistat patients. The need for these assumptions, outlined in Section 5.3 and discussed further in Section 5.11, adds important uncertainty to the conservative comparison to orlistat that the model cannot address. The estimated ICER for NB32 versus orlistat should be interpreted with particular caution; the true ICER could well imply NB32 is a cost-effective alternative to orlistat adjunct therapy, but it is beyond the capability of the economic analysis to demonstrate this

Consequences of conservative assumptions – limitations

1. The analysis is 'blind' to cost and HRQoL benefits of weight reduction for people with known risk factors (possibly over 60 health events – other than just T2DM, MI and stroke) – NB32 is therefore inherently underestimated
2. Natural history model captures weight reduction upon **'time'** to co-morbidity onset rather than weight reduction upon **'probability'** of co-morbidity onset
3. Weight regain is assumed to begin upon treatment discontinuation and treatment is assumed to end at the limit of the clinical data. This underestimates the benefits of NB32 is people continue to gain benefits of weight reductions after treatment discontinuation

Subgroup analysis results (deterministic)

- **People with T2DM at baseline**

Technologies	Total		Incremental		ICER (QALYs)	
	Costs	QALYs	Costs	QALYs	Versus baseline (SM)	Incremental
SM	£10,199	14.37				
Orlistat	£10,496	14.43	£297	0.06	**£5,059**	**£5,059**
NB32	£11,216	14.44	£720	0.01	**£14,797**	**£72,069**

- **People without T2DM at baseline**

Technologies	Total		Incremental		ICER (QALYs)	
	Costs	QALYs	Costs	QALYs	Versus baseline (SM)	Incremental
SM	£3,844	15.73				
Orlisat	£4,077	15.77	£233	0.04	**£6,283**	**£6,283**
NB32	£4,811	15.80	£734	0.03	**£15,339**	**£28,291**

Note, results rounded to 2 decimal places

40

Source: see section 5.9 of the CS.

To note:

all comparisons to orlistat should be interpreted with care, as data regarding comparisons of NB32 to orlistat in people with T2DM are extremely limited as shown, in Section 4.10 of the CS.

Results for NB32 versus standard management in these subgroups are broadly in line with those produced in the model base case (i.e. assuming 33.2% of people with T2DM at baseline).

Company's scenario analysis

Scenario			ICERs NB32 vs	
Model setting	**Base case**	**Scenario tested**	**Orlistat**	**Standard Management**
Base case			**£32,084**	**£13,647**
Weight regain	3 years	2 years	**£41,016**	**£14,113**
Weight regain	3 years	5 years	**£29,739**	**£11,880**
Cost of T2DM	£347.57	£175.86 in Year 1 only	**£36,096**	**£13,764**
Utility model	Tobit	OLS	**£36,771**	**£10,285**
AE costs	All GP	All outpatient	**£36,492**	**£15,130**
Discounting	3.5% for costs & effects	1.5% for costs & effects	**£28,323**	**£9,969**
Time horizon	Lifetime	15 years	**£53,514**	**£22,763**

*Company's base case ICER was £32,084 vs orlistat and £13,647 vs standard management

41

Source: see section 5.8 of the company's submission for full scenario analysis results.

Key areas of uncertainty in the comparison with orlistat

- HRQOL (using OLS model instead of the Tobit model) – increases ICER by 4k if OLS used
- Rate of weight regain – shorter period increases the ICER

One-way sensitivity analysis

The most influential parameters that significantly changed the ICERs were those relating to the Tobit model, the discount rate for QALYs (using 1.5% instead of 3.5%) and adjusting the relative efficacy from the ITC. When comparing NB32 to SM no parameter change raised the ICER more than £20,000 per QALY gained. When comparing NB32 to orlistat there was large variations in the ICER, mostly attributed to the uncertainty in the ITC.

ERG comments on the cost-effectiveness results

Deterministic results

- Company did not run enough patient samples to produce stable ICERs
- ERG estimates that model should run for at least 1,500 samples (company ran 1000) to produce stable results (where convergence occurs), hence results should be interpreted with caution
  - In contrast Ara et al. used a cohort of 1,000,000 patients in their patient-level simulation

Probabilistic results

- Probabilistic sensitivity analysis (PSA) excluded key input parameters (TTD, natural history of BMI model, obesity-related events). Also not explored in deterministic SA
- PSA did not run enough samples to produce convergence and stable results (usually a min of a 1,000 but company ran 500)
- Model not fit for purpose due to extremely long run times and inability to perform appropriate PSA and check the model's internal validity to usual standards
- Probabilistic results are preferred for decision-making (NICE DSU guidance) – if the PSA is flawed so is the estimation of mean outcomes

42

Source: see section 5.2.10 of the ERG report

ERG's amended base case analysis

ERG able to adjust/correct some of the highlighted issues in its base-case:

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/m^2)
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

43

Source: see section 5.3 of the ERG report.

ERG comments

The ERG identified numerous issues; the most important ones are summarised in Table 5.20. Several issues still remain unexplored, some of which were expected to non-conservative so the results should be interpreted with extreme caution. The interpretation and validity of the results are particularly hampered given that the company's model did underestimate TTD, did not incorporate behaviour modification interventions, bariatric surgery and re-treatment nor accurately reflected patients' expected quality of life and costs associated with resource use.

BMI development (i.e. weight regain model) was not accurately reflected in the model (due to lack of an updating event or integration of the BMI function) which could significantly bias the results in favour of NB32.

ERG amended deterministic base case results

Technology	Total		Incremental		ICER (QALYs)	
	Costs	QALYs	Costs	QALYs	Versus baseline (SM)	Incremental
ERG base case 1st run						
SM	£5,964	15.11				
orlistat	£6,275	15.20	£311	0.09	**£3,701**	**£3,701**
NB32	£7,017	15.21	£742	0.01	**£10,510**	**£45,694**
ERG base case 2nd run						
SM	£6,141	14.97				
orlistat	£6,455	15.06	£314	0.09	**£3,466**	**£3,466**
NB32	£7,188	15.08	£733	0.02	**£9,813**	**£38,871**
ERG's replication of the company's base case						
SM	£5,974	15.29				
orlistat	£6,219	15.33	£245	0.04	**£5,865**	**£5,865**
NB32	£6,948	15.36	£729	0.03	**£15,568**	**£34,994**

*Company's base case ICER was £32,084 vs orlistat and £13,647 vs standard management
Note, results rounded to 2 decimal places

44

Source: see section 5.3 of the ERG report.

ERG amended base case

Restricted to a 1,000 patient runs in line with company, given the flaws highlighted

Analysis ran twice (using different random numbers and patient samples) and obtained different results. ERG also re-ran company's base case and obtained different results

ERG comments

The large variations seen in the ICERs with different patient runs is of a particular concern to the ERG. Two runs of the ERG base case lead to the ICER varying by as much as £7,000 per QALY gained. The ERG conclude that this limits the models value for the current decision problem so the result should be interpreted with extreme caution.

ERG additional analyses (conditional on ERG's base case)

Technology	Total		Incremental		ICER (QALYs)	
	Costs	QALYs	Costs	QALYs	Versus baseline (SM)	
Exploratory analysis – using instantaneous weight regain at 3 years						
SM	£6,007	15.09				
orlistat	£6,311	15.17	£304	0.08	**£3,600**	**£3,600**
NB32	£7,048	15.21	£737	0.04	**£10,021**	**£37,947**
Exploratory analysis – lower proportion of people with T2DM (15%)						
SM	£4,702	15.45				
orlistat	£4,992	15.53	£290	0.08	**£3,738**	**£3,738**
NB32	£5,740	15.55	£748	0.02	**£10,013**	**£28,687**
Subgroup analysis – people without T2DM						
SM	£3,565	15.66				
orlistat	£3,844	15.74	£279	0.08	**£3,488**	**£3,488**
NB32	£4,603	15.77	£759	0.03	**£9,594**	**£25,744**
Subgroup analysis –people with T2DM only						
SM	£11,173	13.98				
orlistat	£11,527	14.09	£354	0.10	**£3,435**	**£3,435**
NB32	£12,213	14.08	£686	-0.01	**£10,535**	**Dominated**

Source: see section 5.3.2 of the ERG report.

Conclusion from ERG analysis

The ERG conclude 'the deterministic ERG base-case ICER of NB32 versus orlistat is estimated to range between £38,871 and £45,694 per QALY gained (based on different random numbers and different samples of patients), and the remaining issues/methodological flaws highlighted in the report, uncertainty around the cost effectiveness estimates of NB32 remains substantial'

Innovation & equalities

- Company considers NB32 to be innovative:
  - first oral intervention with a multi-modal mechanism of action that is thought to work through actions in the hypothalamus and the dopaminergic reward system to reduce hunger and reward-driven eating
  - provides a new pharmacological treatment option for a disease of increasing prevalence and substantial burden
  - once people withdraw from current treatment there is a lack of safe and effective pharmacological options in current practice
- Company did not identify any potential equality issues

Key issues: cost effectiveness (1)

1. Model structure

- Model structure does not consider retreatment, behaviour modification, and bariatric surgery. What is the committee's view on the model structure?

2. Model implementation

- The ERG highlighted that the model is very slow to run, large variation in ICERs when different random numbers are used and small number of PSA runs, and BMI updated only when events occur. What is the committee's view on the validity of the model and robustness of the results?

3. Population

- Should the cost-effectiveness be considered in the entire population or in subgroups with/without T2DM? What are the characteristics of the population that should inform the model?

47

Key issues: cost effectiveness (2)

4. Modelling treatment

- What clinical data is appropriate to inform the model? Duration of effect: how fast is weight regained after treatment discontinuation? Treatment duration: is time on treatment appropriately modelled?

5. Utilities

- Is the Public Health England weight management tool appropriate to derive the utilities?

6. Innovation.

- Does the committee consider NB32 to be an innovative therapy?

48

Authors

- **Hamish Lunagaria**
  Technical Lead
- **Zoe Charles**
  Technical Adviser
- with input from the Lead Team – Mohit Mishra, Rita Faria and Pamela Rees

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal

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

Final Scope

Remit/appraisal objective

To appraise the clinical and cost effectiveness of naltrexone-bupropion prolonged release within its licensed indication, in addition to diet and physical activity, for the management of people with obesity or overweight with risk factors.

Background

Overweight and obesity is a chronic condition characterised by increased body fat. People who are overweight or obese are at an increased risk of developing cardiovascular disease, type 2 diabetes, atherosclerosis (the presence of fatty deposits in the arteries), hypertension and dyslipidaemia (abnormal levels of fats in the blood). The most common method for measuring obesity is body mass index (BMI) which is calculated as the ratio of weight to height squared. In adults of European family origin, overweight is typically defined by a BMI of 25 kg/m^2 to <30 kg/m^2 and obesity by a BMI of 30 kg/m^2 or more (an appropriate adjustment of BMI for other ethnic groups is necessary).

In England, 24% of adults are obese and a further 36% are overweight. Of obese adults, seven in ten are Class I obese, with a BMI between 30 and 35. Around one in ten obese adults are morbidly obese, with a BMI above 40[1]. The prevalence of obesity has seen a sharp increase from the 1990s[2]. By 2050 the prevalence of obesity is predicted to affect 60% of adult men, 50% of adult women[2]. Drug items dispensed for managing obesity rose 44 per cent from 2012 to 563,000 in 2013)[3].

Current management of overweight and obesity includes dietary and lifestyle advice, behaviour modification, pharmacological treatments and surgical intervention. Specialist multi-disciplinary weight management interventions (known as tier 3 interventions) are also used in current practice. Tier 3 interventions include dietary, lifestyle and behaviour modification with or without drug therapy. NICE clinical guideline 189 'Obesity: identification, assessment and management' recommends that drug therapy with orlistat should only be considered after dietary, physical activity and behavioural approaches have been started and evaluated. It recommends orlistat for the management of obesity in people with a BMI of 30 kg/m^2 or more, and in people with a BMI of 28 kg/m^2 or more and significant comorbidities. If dietary and lifestyle advice, behaviour modification and drug treatments are unsuccessful, the NICE clinical guideline recommends bariatric surgery for people with: a BMI of 40 kg/m^2 or more; a BMI of between 35 kg/m^2 and

40 kg/m^2 with significant comorbidities, a BMI between 30 kg/m^2 and < 35 kg/m^2 and with recent-onset of type 2 diabetes (surgery can be considered for people of Asian family origin who have recent-onset type 2 diabetes at a lower BMI than other populations).

The technology

Naltrexone-bupropion (Mysimba, Orexigen Therapeutics) is a fixed dose combination of naltrexone and bupropion administered orally in a prolonged-release tablet. Naltrexone is an opioid receptor antagonist and bupropion is a dopamine and noradrenaline reuptake inhibitor. The exact neurochemical appetite suppressant effect of naltrexone-bupropion is not fully understood. It is thought to stimulate pro-opiomelanocortin neuronal firing and modulates food cravings through an effect on the reward pathways of the brain.

Naltrexone-bupropion has marketing authorisation in Europe 'as an adjunct to a reduced-calorie diet and increased physical activity for the management of weight in adults with a BMI of ≥ 30 kg/m^2 (obese) or ≥ 27 kg/m^2 to < 30 kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities'.

Intervention(s)	Naltrexone-bupropion prolonged-release
Population(s)	Adults who have a BMI of; • ≥ 30 kg/m^2 (obese) or • ≥ 27 kg/m^2 to < 30 kg/m^2 (overweight) in the presence of one or more weight- related co-morbidities
Comparators	• Standard management without naltrexone-bupropion • Orlistat (prescription dose)
Outcomes	The outcome measures to be considered include: • BMI • weight loss • percentage body fat • waist circumference • incidence of type 2 diabetes • cardiovascular events • mortality • adverse effects of treatment • health-related quality of life Where information on clinical endpoints is unavailable, consideration may be given to surrogate end-points such

	as: • glycated haemoglobin (HbA1c) • cholesterol levels and lipid profiles (including LDL and HDL) • blood pressure
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.
Other considerations	If the evidence allows, the following subgroup should be considered: people with type 2 diabetes. Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.
Related NICE recommendations and NICE Pathways	**Related Guidelines**: Guideline in development 'Obesity: guidance on the prevention, identification, assessment and management of overweight and obesity in adults and children (update)'. Date of publication November 2014. Clinical Guideline No. 189, ,'Obesity: guidance on the identification, assessment and management obesity in adults and children' Date of publication November 2014 Clinical guideline No. 43. 'Obesity prevention in adults and children' Guidance updated March 2015 **Related Interventional Procedures:** Interventional Procedure Guideline No. 432, November 2013, 'Laparoscopic gastric plication for the treatment of severe obesity'. Interventional Procedure Guidance No. 471, November 2012, 'Implantation of a duodenal-jejunal bypass sleeve for managing obesity'. **Related Public Health Guidance/Guidelines:**

	Public Health Guideline No. 53. 'Weight management: lifestyle services for overweight or obese adults'. Publication date May 2014. Public Health Guideline No. 47, October 2013, 'Managing overweight and obesity among children and young people'. Review proposal date 2017. Public Health Guideline No. 42, November 2012, 'Obesity – working with local communities'. Review proposal date 2017. **Related Quality Standards:** Obesity: clinical assessment and management [QS127] (adults). Published August 2017. Obesity in adults: prevention and lifestyle weight management programmes [QS111]. Published January 2016 Obesity in children and young people: prevention and lifestyle weight management programmes [QS94]. Published July 2015 **Related NICE Pathways:** NICE Pathway: Obesity, Pathway updated August 2016 http://pathways.nice.org.uk/pathways/obesity NICE Pathway: Obesity: working with local communities, Pathway updated: March 2016. http://pathways.nice.org.uk/pathways/obesity-working-with-local-communities
Related National Policy	NHS England (2013) '2013/14 NHS Standard contract for severe and complex obesity (all ages)'. A05/S/a. NHS England (2013) 'Clinical commissioning policy and specialised obesity surgery'. NHS England/A05/P/a.

References

1. Obesity statistics; House of Commons Library. Accessed September 2016.
2. Adult Obesity; Public health England. Accessed September 2016
3. HSCIC Statistics on obesity, physical activity and diet; England 2015. Accessed September 2016

Appendix C

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

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

Matrix of consultees and commentators

Consultees	Commentators (no right to submit or appeal)
Company • Orexigen Therapeutics (naltrexone/bupropion prolonged-release) Patient/carer groups • Beat: Beating eating disorders • BEMDA: Black and Ethnic Minority Diabetes Association • Black Health Agency • Blood Pressure UK • British Cardiac Patients Association • British Obesity Society • Cardiovascular Care Partnership UK • Diabetes Research & Wellness Foundation • Diabetes UK • HEART UK • HOOP UK • InDependent Diabetes Trust • Muslim Health Network • National Obesity Forum • Network of Sikh Organisations • Overeaters Anonymous • South Asian Health Foundation • Specialised Healthcare Alliance • Surya Foundation • Weight Concern Professional groups • Association for the Study of Obesity • Association of British Clinical Diabetologists • British Association for Nursing in Cardiac Care • British Cardiovascular Intervention Society	General • Allied Health Professionals Federation • Board of Community Health Councils in Wales • British Cardiovascular Industry Association • British National Formulary • Care Quality Commission • Department of Health, Social Services and Public Safety for Northern Ireland • Diabetes UK Cymru • Healthcare Improvement Scotland • Medicines and Healthcare products Regulatory Agency • National Association of Primary Care • National Pharmacy Association • NHS Alliance • NHS Commercial Medicines Unit • NHS Confederation • Scottish Medicines Consortium Possible comparator manufacturers • Actavis (orlistat) • Almus pharmaceuticals (orlistat) • GlaxoSmithKline (orlistat, bupropion) • Roche (orlistat) • Teva UK (orlistat) Relevant research groups • British Society for Cardiovascular Research • Cochrane Heart Group • Cochrane Metabolic & Endocrine Disorders Group • Cochrane Peripheral Vascular Disease Group • Cochrane Public Health Group

Consultees	**Commentators (no right to submit or appeal)**
• British Cardiovascular Society • British Dietetic Association • British Geriatrics Society • British Heart Foundation • British Hypertension Society • British Nutrition Foundation • Diabetes Specialist Nurses • Dieticians in Obesity Management • Faculty of Public Health Medicine • National Centre for Eating Disorders • National Diabetes Nurse Consultant Group • National Heart Forum UK • Primary Care Cardiovascular Society • Primary Care Diabetes Society • Royal College of General Practitioners • Royal College of Nursing • Royal College of Pathologists • Royal College of Physicians • Royal Pharmaceutical Society • Royal Society of Medicine • Society for Cardiological Science & Technology • Society for Endocrinology • Society for Vascular Technology • Society of Vascular Nurses • The Nutrition Society • The Obesity Management Association • United Kingdom Clinical Pharmacy Association • Vascular Society <u>Others</u> • Department of Health • NHS Bath and North East Somerset CCG • NHS England • NHS Fylde and Wyre CCG • Welsh Government	• CORDA • MRC Clinical Trials Unit • National Institute for Health Research • Wellcome Trust <u>Associated Public Health Groups</u> • Public Health England • Public Health Wales

Appendix C

NICE is committed to promoting equality, eliminating unlawful discrimination and fostering good relations between people who share a protected characteristic and those who do not. Please let us know if we have missed any important organisations from the lists in the matrix, and which organisations we should include that have a particular focus on relevant equality issues.

PTO FOR DEFINITIONS OF CONSULTEES AND COMMENTATORS

Definitions:

Consultees

Organisations that accept an invitation to participate in the appraisal; the company that markets the technology; national professional organisations; national patient organisations; the Department of Health and the Welsh Government and relevant NHS organisations in England.

The company that markets the technology is invited to make an evidence submission, respond to consultations, nominate clinical specialists and has the right to appeal against the Final Appraisal Determination (FAD).

All non-company consultees are invited to submit a statement[1], respond to consultations, nominate clinical specialists or patient experts and have the right to appeal against the Final Appraisal Determination (FAD).

Commentators

Organisations that engage in the appraisal process but that are not asked to prepare an evidence submission or statement, are able to respond to consultations and they receive the FAD for information only, without right of appeal. These organisations are: companies that market comparator technologies; Healthcare Improvement Scotland; other related research groups where appropriate (for example, the Medical Research Council [MRC], National Cancer Research Institute); other groups (for example, the NHS Confederation, NHS Alliance and NHS Commercial Medicines Unit, and the British National Formulary.

All non-company commentators are invited to nominate clinical specialists or patient experts.

[1]Non-company consultees are invited to submit statements relevant to the group they are representing.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

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

Company evidence submission

January 2017

File name	Version	Contains confidential information	Date
ID757 NB32 company evidence submission redactedv2.docx	**3**	**Yes**	**01.02.2017**

Instructions for companies

This is the template for submission of evidence to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Please note that the information requirements for submissions are summarised in this template; full details of the requirements for pharmaceuticals and devices are in the user guide.

This submission must not be longer than 250 pages, excluding appendices and the pages covered by this template.

Companies making evidence submissions to NICE should also refer to the NICE guide to the methods of technology appraisal and the NICE guide to the processes of technology appraisal.

Contents

Tables

Figures

1 Executive summary

Naltrexone 32mg plus bupropion 360mg (NB32) is an innovative combination therapy that offers a new pharmacological treatment option for the management of weight in adults who are obese, or overweight with one or more weight-related comorbidity. NB32 has a novel, multi-modal mechanism of action (MoA), which is thought to target hypothalamic regions responsible for appetite and energy expenditure and mesolimbic circuits that influence reward pathways to affect eating behaviours. This represents a unique approach to weight loss management.[1]

Overweight and obesity are characterized by excess body weight, typically measured by body mass index (BMI) in clinical practice. In the UK, the rates of obesity (BMI 30–40kg/m^2 or more) have more than doubled in the last 25 years, while the prevalence of overweight (BMI 25–29.9kg/m^2) has remained broadly stable at approximately 40%.[2, 3] The fundamental cause of weight gain is an energy imbalance between calories consumed from food and drink and calories expended through energy expenditure; over time, this results in abnormal or excessive fat accumulation.[4] However, overweight and obesity have a complex aetiology, involving environmental, social and economic factors, and there is growing evidence to suggest that hedonic and as well as homeostatic regulation of food intake play an important role.

From a patient perspective, the physical and mental burden of obesity includes difficulties with physical activity, joint pain and depression that can adversely impact quality of life.[4-9] Overweight and obese patients are also at increased risk of developing Type 2 diabetes mellitus (T2DM) and cardiovascular (CV) disease among other co-morbidities that can further impact patient quality of life, and contribute to a significant economic burden of disease to wider society. A report from 2007 estimated that National Health Service (NHS) costs attributed to elevated BMI were £4.2 billion, with further indirect costs amounting to £15.8 billion.[2]

In NHS England, the initial standard of care is to advise calorie-controlled diets, increased physical activity and behaviour modification. However, many patients do not achieve adequate weight loss with such measures. In these patients, pharmacological treatment that can ameliorate weight-related health risks and improve patient health-related quality of life (HRQL) should be considered. In the UK,

orlistat is the only available pharmacological product for weight management (in conjunction with a mildly hypocaloric diet).[10] Orlistat works by reducing the absorption of dietary fats[11], a mechanism of action that causes significant gastrointestinal side effects, including diarrhoea, anal leakage and increased defecation, particularly in individuals who do not adhere to a low-fat diet.[12] Such adverse events (AE) can severely impact patient quality of life, and often lead to treatment discontinuation. There are also concerns of maintained effectiveness with modest weight loss observed after 1-year of orlistat treatment.[12, 13]

There is a strong unmet medical need for novel, orally effective and well-tolerated pharmacological therapies for patients who do not achieve adequate weight loss through dietary changes and exercise. NB32 offers an alternative pharmacological treatment option with an improved, multi-modal MoA and favourable AE profile to patients who are currently treated with orlistat in clinical practice. More importantly, in the absence of a better treatment option currently available, NB32 offers a pharmacological treatment option with proven weight loss efficacy compared to standard management without NB32 (hereafter referred to as standard management).

In the pivotal trial programme (see Section 1.3), treatment with NB32 resulted in early and sustained weight loss which was significantly greater than that observed with standard management. Patients who completed 56 weeks of NB32 treatment in line with the licensed dosing schedule showed a least squares (LS) mean weight loss of 11.7%, a noteworthy improvement of greater magnitude than that commonly seen in this patient population. Significant improvements were also observed across many cardiometabolic risk parameters that may be associated with a reduced risk of CV events, as well as improvements in some diabetic-specific risk factors. Furthermore, patients treated with NB32 reported significant improvements in disease-specific HRQL, and significant and sustained improvements in control of eating and reduced food cravings. Importantly, NB32 was well tolerated with a transient and readily manageable AE profile.

The economic analysis supporting this submission takes a robust and inherently conservative approach to estimate the cost effectiveness of NB32 adjunct therapy for NHS England patients. The methodology is consistent with a previous high-quality National Institute for Health Research (NIHR)-funded systematic analysis of

competing drug treatments for overweight and obese patients.[14] Even with the implicitly conservative nature of the model, results from the analysis show NB32 to be cost effective as an adjunct to standard management, for patients who would otherwise receive standard management alone.

Robust economic comparison to orlistat was challenging. In addition to familiar challenges inherent in comparing across clinical studies conducted at different times, in different regions, using different lifestyle modification programmes and with different subject discontinuation rates, the comparison was further limited by the inconsistency between published clinical trial designs and clinical practice in light of the regulatory treatment discontinuation rule for orlistat. Using the best available data and most plausible assumptions, and considering the inherently conservative features of the economic analysis, NB32 may too represent a cost-effective alternative to orlistat for NHS England patients.

In conclusion, NB32 offers a tolerable, highly effective and cost-effective pharmacological treatment option for overweight or obese patients with significant disease burden and restricted treatment choice in current clinical practice.

1.1 Statement of decision problem

The decision problem addressed in this appraisal is in line with that described in the final scope issued by NICE in October 2016.[15] The decision problem is presented in Table 1.

Table 1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	Adults who have a BMI of: • ≥30kg/m^2 (obese) or • ≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities	Adults who have a BMI of: • ≥30kg/m^2 (obese) or • ≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities	-
Intervention	Naltrexone-bupropion prolonged-release	Naltrexone-bupropion prolonged-release	-
Comparator (s)	• Standard management without naltrexone-bupropion • Orlistat (prescription dose)	• Standard management without naltrexone-bupropion • Orlistat (prescription dose)	-
Outcomes	• BMI • Weight loss • Percentage body fat • Waist circumference • Incidence of Type 2 diabetes • Cardiovascular events • Mortality • Adverse effects of treatment • Health-related quality of life	• Weight loss • Percentage body fat • Waist circumference • Incidence of Type 2 diabetes • Cardiometabolic parameters • Mortality • Adverse effects of treatment • Health-related quality of life	Key outcomes captured in pivotal trial programme
Economic analysis	• The cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year	• The cost effectiveness of treatments is expressed in terms of incremental cost per quality-adjusted life year	-

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
	• The time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared • Costs will be considered from an NHS and Personal Social Services perspective	• The time horizon for estimating clinical and cost effectiveness reflects the lifetime of patients • Costs are considered from an NHS and Personal Social Services perspective	
Subgroups to be considered	People with Type 2 diabetes	People with Type 2 diabetes; the COR-DM study provides data for this subgroup	-
Special considerations including issues related to equity or equality	None specified	None specified	-

Key: BMI, body mass index; CV, cardiovascular; NB32, naltrexone 32mg plus bupropion; NICE, National Institute for Health and Care Excellence; T2DM, type 2 diabetes mellitus

1.2 Description of the technology being appraised

Details of the technology being appraised in this submission are summarised in Table 2.

Table 2: Technology being appraised

UK approved name	Naltrexone plus bupropion (NB32)
Brand name	Mysimba® (US brand name – Contrave®)
Marketing authorisation/CE mark status	Positive opinion from the CHMP was received on 18 December 2014. Marketing authorisation was received on 26 March 2015.
Indications and any restriction(s) as described in the summary of product characteristics	The indication for NB32 is as follows: *"Mysimba is indicated, as an adjunct to a reduced-calorie diet and increased physical activity, for the management of weight in adult patients (≥18 years) with an initial Body Mass Index (BMI) of* • *≥30kg/m^2 (obese), or* • *≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities (e.g., Type 2 diabetes, dyslipidaemia, or controlled hypertension)"* Treatment with Mysimba should be discontinued after 16 weeks if patients have not lost at least 5% of their initial body weight.
Method of administration and dosage	Mysimba is orally administered. Each tablet contains 8mg naltrexone and 90mg bupropion hydrochloride. Dose should be escalated for the first 4 weeks as follows: • Week 1: One tablet in the morning • Week 2: One tablet in the morning and one tablet in the evening • Week 3: Two tablets in the morning and one tablet in the evening • Week 4 and onwards: Two tablets in the morning and two tablets in the evening
Key: CHMP, Committee for Medicinal Products for Human Use; NB, naltrexone plus bupropion	

1.3 Summary of the clinical effectiveness analysis

A comprehensive clinical trial programme supports the use of NB32 for the management of weight in adult patients who are overweight or obese with one or more weight-related co-morbidities. This clinical trial programme includes four pivotal

randomised controlled trials (RCTs), which provide evidence of the clinical effectiveness of treatment with NB32 as an adjunct to standard management, compared with placebo as an adjunct to standard management.

A summary of the Contrave Obesity Research (COR) trial programme is provided below:

COR-I[16]

- Phase III, multicentre, randomised, double-blind, placebo-controlled, 56-week RCT comparing the clinical efficacy and safety of NB32 versus placebo in adult patients with obesity or overweight and controlled hypertension and/or dyslipidaemia. Standard management in this trial consisted of customary diet and behaviour modification.
- Primary efficacy endpoint analyses in the modified intent-to-treat (mITT) population showed LS mean weight loss was greater for patients treated with NB32 (6.1%) compared to placebo (1.3% [p<0.001]). Also, the proportion of patients with ≥5% weight loss was greater in the NB32 arm (48%) compared to the placebo arm (16% [p<0.001]) at Week 56.
- Secondary efficacy endpoint analyses demonstrated that a significantly greater proportion of patients treated with NB32 had ≥10% weight loss compared to placebo (20% vs 7%, respectively; p<00001). NB32 also resulted in improvements in numerous cardiometabolic parameters, including reductions in waist circumference and triglycerides, and an increase in high-density lipoprotein (HDL) cholesterol.

COR-II[17]

- Phase III, multicentre, randomised, double-blind, placebo-controlled, 56-week RCT comparing the clinical efficacy and safety of NB32 versus placebo in adult patients with obesity or overweight and controlled hypertension and/or dyslipidaemia. Standard management in this trial consisted of customary diet and behaviour modification.
- Primary efficacy endpoint analyses in the mITT population, showed LS mean weight loss was significantly greater for patients treated with NB32 (-6.5%) compared to placebo (-1.9% [p<0.001]) at Week 28. Furthermore, the

proportion of patients with ≥5% weight loss at Week 28 was greater in the NB32 arm (55.6%) compared to the placebo arm (17.5% [p<0.001]).

- Secondary efficacy endpoint analyses showed weight loss was maintained to Week 56 where a statistically significantly greater decrease was observed in the NB32 arm than in the placebo arm (-6.4% vs -1.2%, respectively [p<0.001]). In addition, NB32 was associated with a significantly larger proportion of patients achieving ≥10% and ≥15% weight loss compared to placebo at Weeks 28 and 56. NB32 also resulted in improvements in various cardiometabolic parameters, including reductions in waist circumference and triglycerides, and increased HDL cholesterol.

COR-BMOD[18]

- Phase III, multicentre, randomised, double-blind, placebo-controlled, 56-week RCT comparing the clinical efficacy and safety of NB32 versus placebo in adult patients with obesity or overweight and controlled hypertension and/or dyslipidaemia. Standard management in this trial consisted of intensive behaviour modification.
- Primary efficacy endpoint analyses at Week 56 showed patients in the NB32 arm had significantly greater LS mean percent change in body weight compared to the placebo arm (-9.3 vs -5.1, respectively [p<0.001]). More patients also achieved ≥5% weight loss when treated with NB32 (66.4%) compared to placebo (42.5% [p<0.001]).
- Secondary efficacy endpoint analyses showed that a significantly greater proportion of patients in the NB32 arm achieved ≥10% weight loss (41.5%) compared to the placebo arm (20.2% [p<0.001]). NB32 also resulted in improvements in various cardiometabolic parameters, including reductions in waist circumference and triglycerides and an increase in HDL cholesterol.

COR-DM[19]

- Phase III, multicentre, randomised, double-blind, placebo-controlled, 56-week RCT comparing the clinical efficacy and safety of NB32 versus placebo in adult patients with T2DM and obesity or overweight. Standard management in this trial consisted of customary diet and behaviour modification.

- Primary efficacy endpoint analyses showed patients in the NB32 arm lost significantly more weight than patients treated with placebo (LS mean percent change: 5.0% vs 1.8%, respectively [p<0.001]). In addition, more patients treated with NB32 achieved ≥5% reduction in body weight compared to the placebo arm (44.5% vs 18.9%, respectively [p<0.001]).
- Secondary efficacy endpoint analyses showed NB32 was associated with an improvement in diabetic-specific outcomes, such as greater reduction in HbA1c (-0.6%) compared to placebo-treated patients (-0.1% [p<0.001]) and fewer patients in the NB32 arm required an increase in dose or the addition of another oral anti-diabetes drug. NB32 also had greater improvements in cardiometabolic risk factors throughout the 56 weeks.

A pooled analysis of the four pivotal COR trials was conducted retrospectively, and showed that 85% of NB32-treated patients who had achieved ≥5% weight loss at Week 16 maintained ≥5% weight loss at Week 56[20], leading to the inclusion of a 16-week discontinuation rule in the summary of product characteristics (SmPC). Across the four pivotal trials, the proportion of patients with ≥5% weight loss at Week 16 ranged from 44.9% to 69.9% (mITT population), but all patients continued to receive treatment up to Week 56 (as per the trial protocol). Primary efficacy endpoint analysis across these trials should therefore be viewed as a conservative estimate of the potential effectiveness of NB32 treatment in clinical practice, given that in clinical practice up to 50% of patients may have discontinued from treatment after 16 weeks. This is supported with post-hoc, pooled analysis that shows the mean reduction in bodyweight for NB32-treated patients from the COR trials who had achieved ≥5% weight loss at Week 16 was 11.3% at Week 56, with 55% of these individuals losing ≥10% of their initial body weight.

NB32 treatment also resulted in significant improvements in the Impact of Weight on Quality of Life-Lite (IWQOL-Lite) questionnaire across the three pivotal trials, COR-I, COR-II and COR-BMOD, (p<0.05 vs placebo)[16-18] and numerically greater improvements seen in the COR-DM study. Although statistical significance was not reached for IWQOL-Lite in this study (potentially due to the underlying burden associated with T2DM), a longer-term reduction in weight may alleviate the burden of T2DM, possibly leading to significant improvements in HRQL. NB32 treatment also resulted in improvements in the Control of Eating (COE) questionnaire item #19,

indicating that ability to control eating behaviour and in particular to resist food cravings was increased. HRQL benefits were generally maintained throughout the studies.

Importantly, NB32 was also generally well tolerated with a transient and manageable AE profile, commonly consisting of nausea, constipation, headache and dizziness. Most treatment-emergent adverse events (TEAE) were considered mild or moderate in severity and in most cases, TEAEs did not result in study discontinuation. Serious TEAE rates were <4% across all four pivotal trials. Only one death was observed in patients treated with NB32; however, this was considered unlikely to be related to the study drug. This safety profile was consistently observed across all patient groups, including patients with T2DM. Compared to orlistat, NB32's AE profile indicates a less disabling and incapacitating safety profile suggesting that patients are more likely to maintain their quality of life when receiving NB32 compared to orlistat.

Data from two Phase IIIb RCTs demonstrate maintenance of clinical effectiveness and a safety profile which does not change with longer-term treatment of more than 56 weeks.

An indirect treatment comparison (ITC) using data from the four, pivotal COR trials and 16 trials investigating orlistat was performed to analyse the number of patients achieving at least a 5% reduction in weight and percent change in weight after 56 weeks. In studies investigating patients without T2DM, orlistat was statistically inferior versus NB32 for patients achieving at least 5% reduction in weight (odds ratio [OR]: 0.77 [95% credible interval [21]: 0.61, 0.96]) and percentage weight change (mean difference [MD]: 1.13% [95% CrI: 0.44, 1.80]). Studies investigating patients with T2DM suggested that NB32 has comparable efficacy to orlistat, for achieving at least 5% reduction in weight (OR: 1.09 [95% CrI: 0.63, 1.88]) and for percentage weight change from baseline (MD: 0.21% [-0.87, 1.30]).

1.4 Summary of the cost-effectiveness analysis

A *de novo* economic model was developed for this appraisal. Based on key evidence identified by a systematic review of economic evidence and appraisal of the modelling implications of the economic consequences of weight change for overweight and obese patients, an individual-level, continuous-time modelling

approach was used. The model harnesses key data and assumptions from a key NIHR-funded, systematically informed, economic appraisal published in the Health Technology Assessment journal.[14] The model captures (i) an immediate effect of weight change upon HRQL, and (ii) HRQL and cost implications of weight change through time-to-event delays to T2DM onset, myocardial infarction, stroke and death. By capturing only these downstream benefits of weight reduction, and through other key assumptions consistent with the NIHR-funded analysis, the approach is inherently conservative; especially considering of the 63 obesity-related health risks and complications listed in 2015 European Guidelines for Obesity Management in Adults.[6]

Base case incremental cost-effectiveness results are summarised in Table 3, with pairwise cost-effectiveness results for NB32 versus standard management and orlistat summarised in Table 4 and Table 5, respectively. It was pivotal for the economic comparison to standard management to reflect response-based regulatory treatment discontinuation rules that were not used in key clinical trials. It was therefore necessary to carefully analyse patient-level data from the COR trials and the cardiovascular outcomes trial of NB32 (NB-CVOT) study described in Section 1.3. Similar analysis of key orlistat trials was not possible because no patient-level data were available. Therefore, conservative assumptions were required for the comparison to orlistat. Base case results versus orlistat (ICER £32,084) should as such be interpreted with caution.

Table 3 shows the base case incremental cost-effectiveness ratio (ICER) versus standard management is £13,647 per quality-adjusted life year (QALY) gained. Given the inherently conservative approach, this is testament to the clear value of NB32 adjunct therapy for NHS patients who would otherwise receive only standard non-pharmacological management. Prior treatment with orlistat adjunct therapy, a drug with a different mechanism of action to NB32, should have no effect on the effectiveness of NB32 adjunct therapy. As such, NB32 is a cost-effective alternative to standard management for both (i) patients who would not be given orlistat and (ii) patients who have previously received orlistat.

The base case ICER versus orlistat adjunct therapy, viewed in isolation, suggests that NB32 adjunct therapy for patients who would otherwise receive orlistat adjunct therapy would not be cost effective at the NICE willingness-to-pay threshold of

£20,000 per QALY gained. The estimated patient QALY benefit and incremental cost of NB32 versus orlistat are small (0.0234 QALYs and £750, and the ICER is therefore sensitive. Given the conservative features of the *de novo* model, covered briefly in this section and described throughout Section 5, it is almost certain that the true incremental costs of NB32 have been overestimated, and true incremental benefits underestimated, in this comparison. Even if estimated incremental costs are assumed to be correct, if 0.0142 incremental QALYs are being masked by the conservative limitations of the economic analysis, the true ICER for NB32 versus orlistat is below £20,000 per QALY gained.

Table 3: Incremental cost-effectiveness results

Technologies	Total			Incremental			ICER (QALYs)	
	Costs	LYG	QALYs	Costs	LYG	QALYs	Versus baseline (SM)	Incremental
SM	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4148	£294	0.0383	0.0531	£5,538	£5,538
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£13,647	£32,084
Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NB32, naltrexone/bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.								

Table 4: Pairwise cost-effectiveness results: NB32 versus standard management

Technologies	Total costs	Total LYG	Total QALYs	Incremental costs	Incremental LYG	Incremental QALYs	ICER (QALYs)
SM	£6,519	33.4768	15.3616				
NB32	£7,563	33.5343	15.4381	£1,044	0.0575	0.0765	£13,647
Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NB32, naltrexone/bupropion; QALY, quality-adjusted life year; SM, standard management.							

Table 5: Pairwise cost-effectiveness results: NB32 versus orlistat

Technologies	Total costs	Total LYG	Total QALYs	Incremental costs	Incremental LYG	Incremental QALYs	ICER (QALYs)
ORL	£6,814	33.5151	15.4148				
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£32,084
Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NB32, naltrexone/bupropion; ORL, orlistat; QALY, quality-adjusted life year.							

2 The technology

2.1 Description of the technology

Brand name: Mysimba®

UK approved name: Naltrexone plus bupropion (NB32)

Therapeutic class: Centrally acting anti-obesity product

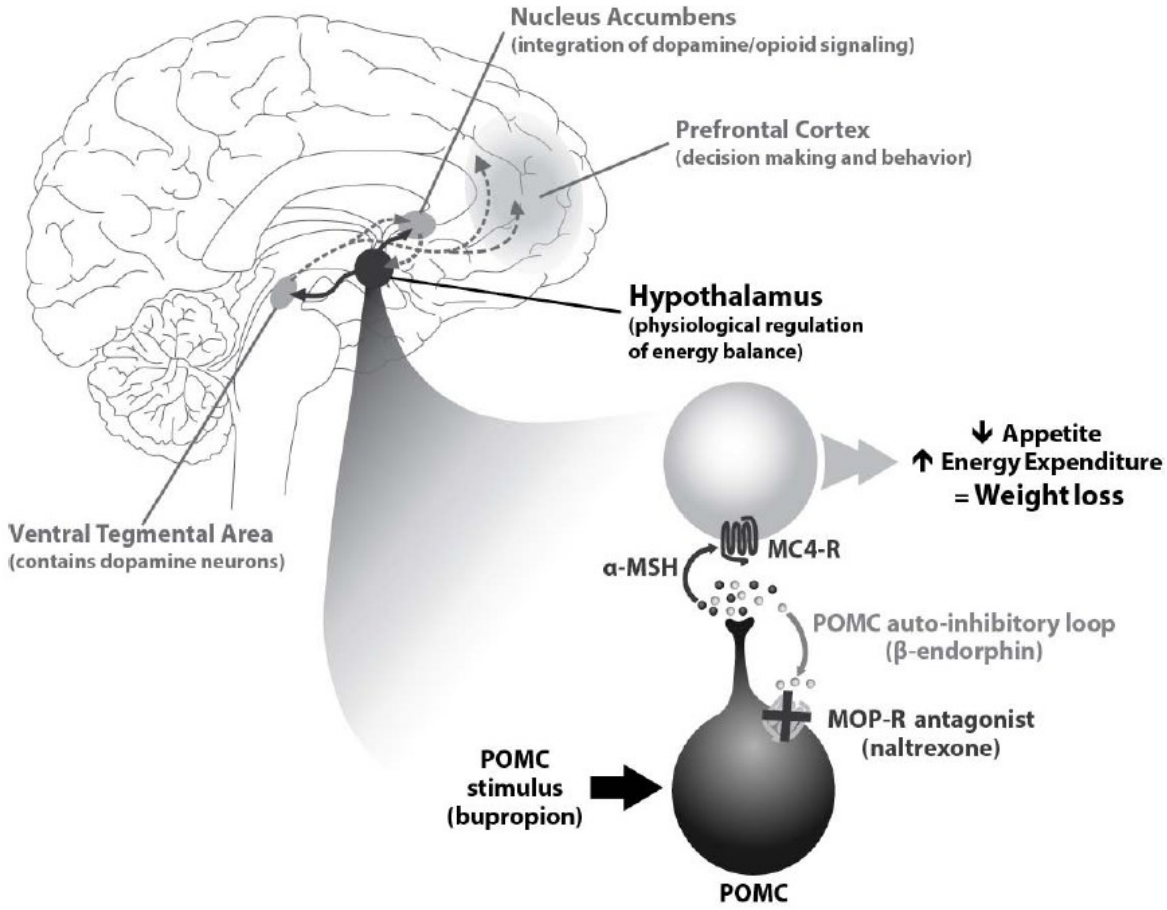
Brief overview of the mechanism of action:

NB32 is a prolonged-release formulation of two currently marketed drugs: naltrexone hydrochloride, a mu-opioid receptor antagonist; together with bupropion hydrochloride, a norepinephrine and dopamine reuptake inhibitor. Both naltrexone and bupropion have been individually used in the European Union (EU) for more than 14 and 25 years, respectively. Naltrexone is approved for opioid dependence[22], and bupropion is indicated for the treatment of major depression and nicotine dependence and has been shown to result in modest weight loss [typically 2–3%] when used for approved indications).[23]

While the exact neurochemical effects of the NB32 combination to reduce food intake are not fully understood, nonclinical studies suggest that the compounds affect complementary pathways in the brain. NB32 acts as a dual pro-opiomelanocortin (POMC) enhancer in the arcuate nucleus of the hypothalamus, initiating a cascade of effects which are thought to result in reduced energy intake and increased energy expenditure.[24, 25] NB32 also targets the mesolimbic reward pathway, which is implicated in the regulation of eating behaviours.[26] By targeting these reward circuits, NB is thought to modulate food craving and reward-driven eating, particularly of highly-palatable foods.[25]

A summary of the mechanism of action of NB32 is presented in Figure 1.

Figure 1: Mechanism of action of naltrexone plus bupropion



Key: MC4-R, melanocortin-4 receptor; MSH, melanocyte-stimulating hormone; MOP-R, μ-opioid receptor; POMC, pro-opiomelanocortin.
Source: Orexigen, 2016[27]

2.2 Marketing authorisation/CE marking and health technology assessment

Mysimba (NB32) received a positive opinion from the CHMP on 18 December 2014 and was approved by the European Medicines Agency (EMA) for a marketing authorisation on 26 March 2015.

The indication for the combination of NB32 is as follows:

"Mysimba is indicated, as an adjunct to a reduced-calorie diet and increased physical activity, for the management of weight in adult patients (≥18 years) with an initial Body Mass Index (BMI) of

- *≥30kg/m^2 (obese), or*

- *≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities (e.g., Type 2 diabetes, dyslipidaemia, or controlled hypertension)*

Treatment with Mysimba should be discontinued after 16 weeks if patients have not lost at least 5% of their initial body weight"

This indication is based on a comprehensive clinical trial programme, including four Phase III studies and a later Phase IIIb study, further discussed in Section 4.

During the assessment for marketing authorisation, the EMA concluded that, overall, the main findings in the pivotal studies were that treatment with NB32 resulted in statistically significant weight loss compared with placebo in adults who are overweight or obese with one or more weight-related comorbidities, including T2DM.[28] Based on retrospective, post-hoc analyses that showed a strong relationship between early and later weight loss (see Section 4.7), a clinically recommended threshold of ≥5% weight loss at Week 16 was incorporated as a discontinuation rule in the summary of product characteristics (SmPC) for Mysimba, as noted above.

The CHMP had some concerns regarding the primary efficacy analysis set of the pivotal Phase III trials, which was the modified intent-to-treat (mITT) population (patients who had at least one post-baseline weight measurement obtained while the patient was still taking study medication) with missing data imputed using the last observation carried forward (LOCF) method (see Section 4.4). This was because a high drop-out rate was observed (see Section 4.5) that included patients who discontinued medication prior to the first post-baseline visit. Pre-specified sensitivity analyses included in the pivotal trials included several analyses using different population sets and different methods of imputation (see Section 4.4). Importantly, all sensitivity analyses substantiated the results of the primary efficacy analysis. Data for the primary efficacy analysis set is presented in Section 4 with sensitivity analyses presented in an appendix, but it should be noted that following discussions with the regulatory authorities, the SmPC presents data for the more conservative analysis set of the intent-to-treat (ITT) population (patients who had at least one post-baseline weight measurement) for mean weight loss data, and the more

conservative imputation method of baseline observation carried forward (BOCF) for responders analysis.

In addition, the CHMP had concerns over the cardiovascular (CV) safety of NB32 due to the sympathomimetic effects of bupropion, and some CV parameter data from the clinical trial programme. Interim data from a Phase IIIb trial (NB-CVOT) that was primarily designed to investigate the CV safety of NB32 in weight management (at the request of the Food and Drug Administration [FDA]), was provided in response to these concerns. Data from this trial did not indicate an increased risk of major adverse cardiovascular events (MACE), and the results were considered to be reassuring with regard to the short- and intermediate-term CV safety. Marketing authorisation was granted on this basis, but a further Phase IV study to assess the effect of NB32 on the occurrence of MACE in overweight and obese patients was requested. Data from this trial are due in 2022. The CHMP also requested additional assessment of the pharmacokinetics of NB32 in patients with renal impairment and in patients with hepatic impairment, as the submitted trials did not collect such data, nor did the Phase III programme allow a direct evaluation of safety in these patient groups. Such a trial is ongoing.

Patients with end-stage renal failure or severe renal or hepatic impairment are listed as a contraindicated patient population in the Mysimba SmPC. Additional contraindicated patient populations include those with uncontrolled hypertension, a current seizure disorder or a history of seizures, and patients with hypersensitivity to any of the excipients. NB32 should also not be administered to patients receiving chronic opiate therapy. In patients requiring intermediate opiate treatment, NB32 should be temporarily discontinued, and the opiate dose should not be increased above the standard dose.

The European Public Assessment Report (EPAR) and the SmPC for Mysimba are provided in Appendix 1.

In addition to European approval, NB32 received marketing authorisation in the United States (US) on 10 September 2014. With regard to further UK HTA, a submission to the Scottish Medicines Consortium (SMC) is planned.

2.3 Administration and costs of the technology

Administration and costs associated with NB32 are summarised in Table 6.

Table 6: Costs of the technology being appraised

	Cost	Source
Pharmaceutical formulation	Prolonged-release tablet	SmPC[29]
Acquisition cost (excluding VAT)[a]	£73.00 per pack of 112 tablets	List price submitted to the Department of Health
Method of administration	Oral administration	SmPC[29]
Doses	Each tablet contains 8mg naltrexone and 90mg bupropion hydrochloride	SmPC[29]
Dosing frequency	Week 1: One tablet in the morning Week 2: One tablet in the morning and one tablet in the evening Week 3: Two tablets in the morning and one tablet in the evening Week 4 and onwards: Two tablets in the morning and two tablets in the evening	SmPC[29]
Average length of a course of treatment	At 16 weeks, treatment should be discontinued if patients have not lost at least 5% of their initial body weight. For patients continuing treatment post 16 weeks, treatment should be continued as long as clinical benefit is observed.	SmPC[29]
Average cost of a course of treatment	Over a lifetime perspective, the economic model predicts total treatment acquisition cost to be £995 for the average NHS patient	Section 5.7.3
Anticipated average interval between courses of treatments	Retreatment with NB32 is not routinely anticipated and thus not modelled.	-
Anticipated number of repeat courses of treatments	Retreatment with NB32 is not routinely anticipated and thus not modelled.	-
Dose adjustments	Outside of the initial 4-week dose escalation, dose escalation or reduction is not recommended.	SmPC[29]
Anticipated care setting	Home setting	SmPC[29]

	Cost	Source
Key: NB32, naltrexone 32mg plus bupropion; NHS, National Health Service; SmPC, summary of product characteristics; VAT, value added tax **Notes:** [a], indicate whether this acquisition cost is list price or includes an approved patient access scheme. When the marketing authorisation or anticipated marketing authorisation recommends the intervention in combination with other treatments, the acquisition cost of each intervention should be presented.		

2.4 Changes in service provision and management

In accordance with the SmPC, NB32 is an oral treatment that can be taken by the patient at home and does not require supervision by a healthcare professional. NB32 should be prescribed as an adjunct to a reduced-calorie diet and increased physical activity; these are standard management approaches currently used in NHS England. As such, no additional infrastructure or NHS resource use is required for the administration of this treatment.

In the first instance, NB32 will be prescribed by a general practitioner (GP). After this, the need for continued treatment should be evaluated at 16 weeks to assess response to therapy and re-evaluated annually. In the UK, monitoring is anticipated throughout treatment with a GP at 3, 6 and 12 months and a practice nurse for the intervening months in line with current practice. This includes annual blood tests to monitor blood glucose levels. Staff and administration costs for the initial prescription, monthly monitoring and annual evaluations are fully accounted for in the economic modelling (see Section 5).

2.5 Innovation

NB32 provides a new pharmacological treatment option for a disease of increasing prevalence and substantial burden (see Section 3.1). Withdrawal of former treatment options due to safety concerns has resulted in a lack of safe and effective pharmacological options in current practice, such that there is currently only one pharmacological treatment (orlistat [Xenical®]) recommended for weight management in NHS England (see Section 3.3).

NB32 is the first oral intervention with a multi-modal mechanism of action that is thought to work through actions in the hypothalamus and the dopaminergic reward system to reduce hunger and reward-driven eating (see Section 2.1). This is an

innovative advancement to the more conventional mechanism of blocking fat absorption, as associated with orlistat (see Section 3.6). In addition, NB32 provides a transient and readily manageable safety profile that is considered favourable to that of orlistat, which can markedly impact patient quality of life (see Section 3.6). During consultation in the pre-scoping stage, the Royal College of Physicians (RCP) stated that[30]:

"overall this combination treatment seems to offer better weight loss than orlistat and with a different adverse effect profile. It therefore represents a new option for patients with obesity"

In addition, when asked if they considered that use of the technology could result in any potential and substantial health-related benefits that are unlikely to be included in the QALY calculation, Diabetes UK stated[30]:

"Yes, more support to people to help them lose weight is welcome. This will go a long way in reducing the rising numbers of Type 2 diabetes and other obesity-related conditions"

Although the weight loss and health-related quality of life (HRQL) benefits assessed as part of the clinical trial programme (see Section 4) should be captured as part of the quality-adjusted life years (QALY) calculation, there is a paucity of evidence to link weight reduction to many important health consequences. As such, it is difficult to capture the full benefit of effective treatment; the RCP specifically identified reductions in conditions such as sleep apnoea, pain from arthritis and low mood as being only partly captured by changes in the QALY measures.[30] This was further validated by clinical consultation.[31] In addition, the weight-loss benefits of NB32 are thought to go a long way in reducing the rising numbers of T2DM and other obesity-related conditions, all of which may not be adequately captured in the QALY.[30]

Furthermore, although not adequately captured in the trial data, it is reasonable to assume that a further long-term benefit of weight loss via pharmacological treatment is prevention of bariatric surgery, a highly burdensome procedure to both patients and the NHS.

In conclusion, the introduction of additional pharmacological treatment options such as NB32 can result in wider-reaching benefits that should be considered alongside the clinical- and cost-effectiveness case based on trial data.

3 Health condition and position of the technology in the treatment pathway

3.1 Disease background

Overweight and obesity imply a condition of excess body weight of a person. In clinical practice, body fatness is generally assessed by the BMI, calculated as body weight (kg) divided by height squared (m^2). The BMI range for normal weight is 18.5–24.9kg/m^2; overweight is 25–29.9kg/m^2; obese is 30–40kg/m^2 and morbidly obese is defined as >40kg/m^2.[3]

In the UK, the rates of obesity have more than doubled in the last 25 years, while the prevalence of overweight has remained broadly stable at approximately 40%.[2] Due to the rapid increase in the prevalence of obesity over the past 30 years, obesity is now recognised as the most prevalent metabolic disease worldwide[32], reaching epidemic proportions in both developed and developing countries and affecting not only adults, but also children and adolescents.[4]

Overall, men are more likely to be overweight; however, women are more likely to be obese.[33] Furthermore, those aged 55–64 years are the most likely to be obese, while 16–24 year olds are least likely.[33] When using BMI as a measure, findings suggest that compared to the general population, obesity prevalence is lower among men from Black African, Indian, Pakistani, and, most markedly, Bangladeshi and Chinese communities. Among women, obesity prevalence appears to be higher for those from Black African, Black Caribbean and Pakistani groups than for women in the general population and lower for women from the Chinese ethnic group.[34]

For both overweight and obesity, the fundamental cause is an energy imbalance between calories consumed from food and drink and calories expended through exercise and energy expenditure; over time, this imbalance results in abnormal or excessive fat accumulation.[4] Globally, the most influential factor for increasing prevalence is the increase in availability of high-fat foods alongside the increasingly sedentary nature of occupations.[4, 35] However, a number of other factors markedly contribute with over 100 difference determinants or variables estimated to directly or indirectly influence energy balance.[2] These include environmental factors, such as development of habits and convenient access to unhealthy foods; economic factors,

such as the typically cheaper cost of processed foods; and biological factors, such as genetic and epigenetic influences (including those affecting the melanocortin system), maternal conditions, use of several medications, and GI microbiome composition.[2, 36]

This increased understanding of the complex aetiology of overweight and obesity has changed the perception that such conditions can be eradicated if people simply ate less and did more[2, 4], and it is now recognised that hedonic and as well as homeostatic regulation of food intake play an important role.

3.2 Effect of disease on patients, carers and society

Overweight and obesity are both associated with a wide range of and debilitating health problems.[5, 6] Of note, available literature focuses on health problems associated with obesity; however, because many people who are overweight will become obese in their lifetime, it is reasonable to assume the comorbidities listed are relevant to both populations.

Co-morbidities associated with increased bodyweight are presented in Table 7. In summary, overweight and obese patients are at higher risk of T2DM and CV disease, as well as many other comorbidities that contribute to a significant economic burden (discussed further below).

Table 7: Comorbidities associated with increased body fatness

Comorbidity	Associated risk
T2DM	• 90% of patients with T2DM have a BMI >23kg/m^2 • Risk of developing T2DM is about 20 times more likely for people who are obese compared to lean people • Increasingly prevalent with over 4 million people being diagnosed with T2DM in the UK in 2016
Hypertension	• 5-fold risk in obesity • 66% of hypertension is linked to excess weight
Heart disease	• 35% of ischaemic heart disease is linked to excess weight • Obesity is a contributing factor to cardiac failure in >10% of patients
Dyslipidaemia	• This progressively develops as BMI increases above 21kg/m^2 with a rise in small particle LDL

Comorbidity	Associated risk
Coronary artery disease and stroke	• 2.4-fold increase in obese women and 2-fold increase in obese men under the age of 50 • 70% of obese women with hypertension have left ventricular hypertrophy • Obesity is a contributing factor to cardiac failure in >10% of patients • Overweight/obesity plus hypertension is associated with increased risk of ischaemic stroke
Respiratory effects	• Neck circumference of >43cm in men and >40.5cm in women is associated with obstructive sleep apnoea, daytime somnolence and development of pulmonary hypertension
Cancers	• 10% of all cancer deaths among non-smokers are related to obesity o This increases to 30% for endometrial cancers
Reproductive function	• 6% of primary infertility in women is attributable to obesity • Impotency and infertility are frequently associated with obesity in men
Osteoarthritis	• Frequent association in the elderly with increasing body weight o The risk of disability attributable to osteoarthritis is equal to heart disease and greater than any other medical disorder of the elderly
Key: BMI, body mass index; LDL, low-density lipoprotein; T2DM, Type 2 diabetes mellitus. **Source:** Butland et al. 2007[2]; Diabetes UK[37]; Haslam et al. 2015[5]; Yumuk et al. 2015[6]	

From a patient perspective, the physical burden of excess weight includes factors such as limited mobility, difficulties with physical activity, joint problems, pain and discomfort.[4-8] Overweight and obesity also have a substantial mental health burden and can be associated with sleep apnoea and severe depression.[9, 38] One study of HRQL in patients with chronic health conditions found that clinical depression was highest in participants with a BMI >35kg/m^2.[39] This is supported by a Swedish study that found clinically significant depression to be up to four times higher in severely obese individuals than in similar non-obese individuals.[40] It is this physical and mental health burden that results in patients seeking medical help, rather than the comorbidity risks of being overweight, as these are the factors that often inhibit activities of daily living and adversely impact patient quality of life.

The economic burden of managing overweight and obese patients is substantial, both for the health service and wider society. Due to the wide range of comorbidities,

the medical costs of excess bodyweight vary and are therefore difficult to predict accurately. A report from 2007 estimated that NHS costs attributed to elevated BMI were £4.2 billion, with indirect costs amounting to £15.8 billion.[2] This was expected to rise to £6.3 billion in 2015, £8.3 billion in 2025 and £9.7 billion in 2050.[2, 33] With regard to indirect costs, the estimated economic burden associated with T2DM was £8.8 billion for treatment, intervention and complications of diabetes in 2010–2011.[41] An additional economic burden is related to lost productivity, with obese individuals often taking more short- and long-term sickness absence, due to a range of issues including back problems and sleep apnoea, than workers of a healthy weight.[38]

There are also resource implications for social care services as a result of impaired activity due to excess weight, including housing adaptations, specialised carers trained in manual handling of severely obese people and the provision of appropriate transport and facilities.[42] Despite a paucity of data on caregiver burden, it is reasonable to assume that informal provision of supportive care can also negatively impact the quality of life of family and friends of adults who are overweight or obese.

As such, overweight and obesity not only threatens the health and well-being of individuals, it also places an intolerable burden on society in terms of healthcare costs, on employers through lost productivity and on families because of the increasing burden of long-term chronic disability.[2]

3.3 Clinical pathway of care

In NHS England, weight management options are available across a range of tiered services. While definitions may vary locally, Tier 1 comprises universal services such as health promotion, Tier 2 covers lifestyle interventions, Tier 3 covers specialist weight management services, and Tier 4 covers bariatric surgery.[43] Treatment is based upon a patient's BMI and what, if any, comorbidities are present, as outlined in Table 8.

Table 8: Summary of treatment options for overweight and obese patients

BMI classification (kg/m^2)	**Waist circumference[a]**			**Comorbidities present**
	Low	High	Very high	
Overweight (25–29.9)	1	2	2	3
Obesity I (30–34.9)	2	2	2	3
Obesity II (35–39.9)	3	3	3	4

Obesity III (40 or more)	4	4	4	4
Treatment options				
1	General advice on health weight and lifestyle			
2	Diet and physical activity			
3	Diet and physical activity; consider drugs			
4	Diet and physical activity; consider drugs; consider surgery			

Key: BMI, body mass index.
Notes: [a], for men, waist circumference of less than 94cm is low, 94–102cm is high and more than 102cm is very high. For women, waist circumference of less than 80cm is low, 80–88cm is high and more than 88cm is very high.
Source: NICE, 2014[3]

Most patients have already tried dieting and exercise several times before deciding to seek therapy for weight loss. For those who do seek treatment, in NHS England, the initial standard of care is to advise lower-energy diets, increased physical activity and behaviour modification. The exact nature of these treatments can vary in both style and intensity throughout NHS England and may be delivered by either dieticians, GPs or WeightWatchers®. For patients who have not achieved adequate weight loss (who have not reached their target weight loss, or who have reached a plateau) on such standard management, pharmacological treatment should be considered. Such pharmacological treatments can theoretically help patients maintain compliance with diet and exercise regimens, ameliorate obesity-related health risks and improve quality of life in cases where standard management alone does not suffice.[6]

Currently in the EU, orlistat is the only available, orally effective, pharmacological product for weight management on the market; this is especially problematic given the complex aetiology of the disease across individuals (see Section 3.1). Orlistat is licensed for the treatment of obese patients with a BMI ≥30kg/m^2, or overweight patients (BMI ≥28kg/m^2) with associated risk factors, in conjunction with a mildly hypocaloric diet.[10] Due to its mechanism of action, orlistat is associated with several limitations, as detailed in Section 3.6. Therefore, the potential benefits of the addition of pharmacotherapy to standard management are not generally observed, as use of orlistat remains low.

For patients who have tried both standard management and pharmacological treatment but have not achieved or maintained adequate, clinically beneficial weight

loss, bariatric surgery may be considered. Surgery is only indicated for patients with a BMI ≥40kg/m^2 or between 35kg/m^2 and 40kg/m^2 with other significant disease, and who have failed all non-surgical measures, including intensive management in a Tier 3 service.[3] Therefore, surgery should be considered a last resort for patients who have exhausted all other treatment options, as seen by the limited number of surgeries conducted each year, and is therefore not considered an appropriate comparator to NB32, in line with the final scope for this submission.

NB32 offers a well-tolerated pharmacological treatment option with a novel mechanism of action demonstrated to induce and sustain weight loss. NB32 can be used as an alternative first-line pharmacological treatment in patients for whom orlistat is contraindicated or is not utilised due to physician/patient choice, and patients who persevere with standard management despite the expected lack of effectiveness. NB32 should also be considered for patients who have not achieved adequate weight loss with orlistat treatment, or who did not comply with dietary requirements associated with orlistat, or were unable to tolerate orlistat treatment and who would otherwise revisit standard management measures.

3.4 Life expectancy and patient population

Based on the 2014 Health Survey for England, a total of 11,126,000 adults (aged ≥16) were obese (BMI ≥30kg/m^2). In addition 15,825,000 adults are overweight[44] with around 30% or 4,747,500 having a BMI ≥27kg/m^2. Of these, an estimated 16% will have one or more weight-related comorbidity, equivalent to 779,680 patients. Therefore, a total of 11,905,680 adults in England are overweight or obese with one or more weight-related comorbidities.

Overweight/obesity is the fifth leading risk for global deaths. At least 2.8 million adults die each year as a result of being overweight or obese.[45] In 2004, research by a House of Commons Select Committee estimated that 34,100 deaths were attributable to obesity. This equates to 6.8% of all deaths in England.[46]

3.5 Relevant NICE guidance and clinical guidelines

A summary of relevant NICE guidance and clinical guidelines is presented in Table 9.

Table 9: Relevant NICE guidance and clinical guidelines

Organisation	Title	Date	Summary
NICE Guidance			
NICE CG189[3]	Obesity: identification, assessment and management[a]	2014	• Specialist settings for treating severely obese patients should be equipped with, for example, special seating and adequate weighing equipment • Planned weight management should be tailored to the patient's preferences, initial fitness, health status and lifestyle • Regular, non-discriminatory, long-term follow-up by a trained professional should be offered • BMI should be used as a practical estimate of adiposity in adults o BMI should be interpreted with caution; waist circumference may be used in addition for patients with BMI <35kg/m^2 o Bioimpedance should not be used • BMI should be interpreted with caution in muscular adults o Other populations, such as Asians and older patients, have comorbidity risk factors that are of concern at different BMIs • Assessment of health risks associated with being overweight or obese should be based on BMI and waist circumference • Referral to Tier 3 services should be considered if: o The underlying causes of obesity/overweight need to be addressed; the patient has complex disease states or needs not adequately managed in Tier 2; unsuccessful conventional treatment; drug treatment is being considered for BMI >50kg/m^2; specialist interventions; surgery is being considered • Multi-component lifestyle interventions are the treatment of choice and should include behaviour change strategies • Patients should be encouraged to increase physical activity even if they do not lose weight as a result, because of the other health benefits it can bring • Dietary changes should be tailored to food preferences and unduly restrictive or nutritionally unbalanced diets should not be used

Organisation	Title	Date	Summary
			○ Diets that have a 600kcal/day deficit in combination with expert support and intensive follow-up are recommended for sustainable weight loss ○ Patients should be encouraged to eat a balanced diet in the long term • Pharmacological treatment should be considered only after dietary, exercise and behaviour modification counselling have failed ○ Drug treatment should be monitored regularly and withdrawn if patients do not reach target goals; weight loss may be slower in T2DM patients ○ Orlistat may be prescribed according to license but should not be continued beyond 3 months if patients have not lost ≥5% of their initial body weight
NICE CG43[47]	Obesity prevention	2006	• Managers and health professionals in all primary care settings should ensure that preventing and managing obesity is a priority at both strategic and delivery levels. Dedicated resources should be allocated for action • Interventions to increase physical activity should focus on activities that easily fit into people's everyday life, while dietary interventions should be multicomponent • All community programmes should address the concerns of local people; health professionals should work with shops, supermarkets, restaurants, cafes and voluntary community services to promote healthy eating choices • Health professionals such as occupational health staff and public health practitioners should establish partnerships with local businesses and support the implementation of workplace programmes to prevent and manage obesity • Local authorities should provide tailored advice such as personalised travel plans to increase active travel among people who are motivated to change • Community-based interventions should include awareness-raising promotional activities, but these should be part of a longer-term, multicomponent intervention rather than one-off activities • All workplaces should address the prevention and management of obesity ○ They should provide opportunities for staff to eat a healthy diet and be more physically active; incentives such as contribution to gym membership should be sustained

Organisation	Title	Date	Summary
			○ Workplaces providing health checks should ensure that they address weight, diet and activity
NICE PH53[43]	Weight management: lifestyle services for overweight or obese adults	2014	• An integrated approach to preventing and managing obesity should be adopted; patients should be referred to, or allowed to receive support from (or across) the different service tiers of an obesity pathway, as necessary • Be aware of the effort needed to lose weight, prevent weight regain or avoid any further weight gain; be aware of the stigma of overweight or obese adults and ensure equipment and facilities met the needs of most adults • Raise awareness of local weight management issues among commissioners, health and social care professionals and the local population • Overweight and obese adults should be referred to a lifestyle weight management programme after measuring a patients BMI and waist circumference • Address the expectations and information needs of adults thinking about joining a lifestyle weight management programme and discuss the importance and wider benefits of making gradual, long-term changes to their dietary habits and physical activity levels • Improve programme uptake, adherence and outcomes by exploring any issues that may affect their likelihood of benefitting from the programme • Commission programmes that include the core components for effective weight loss and that prevent weight gain including dietary intake, physical activity levels and behaviour change • Improve information sharing for people who attend weight management programmes and monitor and evaluate the programmes
NICE QS127[48]	Obesity: clinical assessment and management	2016	• Patients should be informed of their BMI when it is calculated and advised about associated health risks • Adults with BMI ≥30kg/m^2 for whom tier 2 interventions have been unsuccessful should discuss the choice of alternative treatments, including tier 3 services • Adults with a BMI ≥35kg/m^2 who have been diagnosed with T2DM are offered expedited referral for bariatric surgery assessment

Organisation	Title	Date	Summary
			• Adults with a BMI ≥50kg/m^2 are offered referral for bariatric surgery assessment
NICE QS111[49]	Obesity in adults: prevention and lifestyle weight management programmes	2016	• Adults using vending machines in local authority and NHS venues can buy healthy food and drink options • Adults see details of nutritional information on menus at local authority and NHS venues • Adults see healthy food and drink choices displayed prominently in local authority and NHS venues • Adults have access to a publicly available, up-to-date list of local lifestyle weight management programmes • Adults can access data on attendance, outcomes and views of participants and staff from locally commissioned lifestyle weight management programmes • Adults identified as being overweight or obese are given information about local lifestyle weight management programmes • Adults identified as being overweight or obese, with comorbidities, are offered a referral to a lifestyle weight management programme • Adults about to complete a lifestyle weight management programme agree a plan to prevent weigh regain
European Guidelines			
EASO[6]	European guidelines for obesity management in adults	2015	• A comprehensive history, physical examination and laboratory assessment relative to the patient's obesity should be obtained • Although waist circumference can be used as a proxy for abdominal fat, the development of devices and equipment to more accurately measure body fat offer further options outside of BMI • Appropriate goals of weight management emphasise realistic weight loss (generally X to Y% of initial body weight) to achieve a reduction in health risks and should include promotion of weight loss, maintenance and prevention of weight gain • Obesity management should not only focus on BMI reduction; more attention should be paid to waist circumference and the improvement in body composition

Organisation	Title	Date	Summary
			○ Management of comorbidities, improving QoL and well-being of obese patients are also included in treatment aims • In overweight patients (BMI 25–29.9kg/m^2) without overt comorbidities, prevention of further weight gain rather than weight loss may be an appropriate target • A 5–15% weight loss over a period of 6 months is realistic and of proven health benefit; a greater (20% or more) weight loss may be considered for those with greater degrees of obesity (BMI ≥35kg/m^2) • Referral to an obesity specialist should be considered if the patient fails to lose weight in response to the prescribed intervention • General nutrition and dietary advice should include: decrease energy density of food and drinks; decrease portion size; avoid snacking; do not skip breakfast; reduce episodes of loss of control or binge eating • At least 150 min/week of moderate aerobic exercise should be combined with three weekly sessions of resistance exercise to increase muscle strength • CBT elements should form part of routine dietary management or, as a structured programme, form the basis of specialist intervention • Pharmacological treatment should be considered and evaluated after the first 3 months; if >5% weight loss treatment should be continued: ○ Pharmacological options include orlistat, lorcaserin, phentermine/topiramate, bupropion/naltrexone and liraglutide[b] • Surgery is the most effective treatment for morbid obesity in terms of long-term weight loss, improvements of comorbidities and quality of life and decreases of overall mortality.
Key: BMI, body mass index; CBT, cognitive behavioural therapy; NHS, National Health Service; QoL, quality of life; T2DM, Type 2 diabetes mellitus. **Notes:** [a], currently undergoing a review and update; [b], of these, only orlistat is currently available in the UK.			

3.6 Issues relating to clinical practice

For patients who have not achieved adequate weight loss through diet and exercise, and for whom orlistat treatment is unsuitable or ineffective, there is a lack of safe and effective alternative pharmacological options in current practice.

Orlistat, currently the only orally effective pharmacological product available through NHS England, works by reducing the absorption of dietary fats, leaving unabsorbed lipids to be excreted in the faeces.[11] Due to this mechanism of action, orlistat has been associated with significant gastrointestinal (GI) side effects, including diarrhoea, anal leakage and increased defecation, particularly in individuals who do not adhere to a low-fat diet.[12] Such adverse events (AEs) can adversely impact patient quality of life. Furthermore, it has been suggested that, after 1 year on treatment, weight loss with orlistat is modest (around 3% greater than placebo)[12, 13] and plateaus after four months of therapy[1] with weight regain seen after a year on therapy.[50] In addition, orlistat can decrease absorption of fat-soluble vitamins (Vitamins A, D, E and K), and therefore, some patients need multivitamin supplementation.[12] Orlistat may also not be suitable for treatment of patients who have T2DM as the associated dietary requirements required may not be conducive to good glycaemic control.[51]

As a result of the limitations of orlistat and the absence of alternative treatment choice in current practice, some patients are restricted to perseverance with standard management, despite the lack of effectiveness. Therefore, there is a clear unmet medical need for additional, effective and well-tolerated pharmacological therapies that can induce and sustain weight loss in patients who have not achieved and/or sustained adequate weight loss through dietary and exercise changes.

NB32 addresses this unmet medical need. It provides an alternative pharmacological treatment option with an improved, multi-modal mechanism of action and a favourable AE profile to patients treated with orlistat in current clinical practice. More importantly, NB32 offers a pharmacological treatment option with proven weight loss efficacy to patients on standard management in the absence of a better treatment option in current clinical practice.

3.7 Equality

No equality issues related to the use of NB in adults who are overweight or obese have been identified or are foreseen.

4 Clinical effectiveness

4.1 Identification and selection of relevant studies

Search strategy

A systematic literature review (SLR) designed to identify studies of NB32 and potential comparator therapies to treat adults who are overweight or obese with one or more weight-related comorbidities was initiated on 30 May 2016.

Searches were performed in the following electronic databases:

- MEDLINE and MEDLINE-In-Process
- Embase
- The Cochrane Library, including the following
    - Cochrane Central Register of Controlled Trials (CENTRAL)
    - The Cochrane Database of Systematic Reviews (CDSR)
    - Database of Abstracts of Reviews of Effectiveness (DARE)

In addition to the database searches, supplementary searches were also conducted for the following conferences for the past two years:

- International Congress on Obesity (ICO)
- European Congress on Obesity by the European Association for the Study of Obesity (ECO)
- American Diabetes Association (ADA)
- International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual European Congress
- ISPOR Annual International Congress

Reference lists of existing SLRs and meta-analyses identified through systematic searches were hand-searched to identify additional studies.

The search strategies used for clinical effectiveness searches are provided in Appendix 2.

Study selection

The full eligibility criteria applied to the identified evidence base is presented in Table 10. Briefly, clinical trials that investigated adults who are obese, or overweight with

one or more weight-related comorbidities, were included. Interventions of interest included NB32 and orlistat, while comparator therapies included standard management and any other pharmacological treatments for obesity or weight management. Study duration was limited to >1 year total randomised phase to reflect clinical practice.

Table 10: Eligibility criteria for trials to be included in the systematic review

Criteria	Inclusion criteria	Exclusion criteria
Population	Adults who are obese (BMI ≥30kg/m^2) or overweight, according to one of the following definitions: • 25kg/m^2 to 29.9kg/m^2 • ≥27kg/m^2 to <30kg/m^2 • >28kg/m^2 with one or more weight-related comorbidity (T2DM, dyslipidaemia and/or controlled hypertension)	Healthy volunteers Children (age <18 years) Diseases other than that specified in inclusion criteria
Study design	RCTs Non-RCTs Systematic reviews and meta-analyses of RCTs[a]	*In vitro* studies Preclinical studies Comments, letters, editorials Case reports, case series Non-systematic reviews Observational studies
Intervention	Studies assessing at least one of the following interventions will be included: Naltrexone-bupropion Orlistat	Studies that do not assess at least one of the included interventions will be excluded
Comparator	Comparator therapies may include one of the following: Behavioural interventions Lifestyle or dietary modifications Any treatment listed under the interventions Any other pharmacological treatments for obesity or weight management	Studies will not be excluded on comparator therapy if it includes at least one of the treatments listed under the interventions
Study duration	All trials with total randomised phase duration >1 year are included	Studies with <1-year duration
Language	Studies published in English were included Studies published in non-English languages were flagged	Studies will not be excluded on the basis of publication language

Criteria	Inclusion criteria	Exclusion criteria
Key: BMI, body mass index; RCT, randomised controlled trial; T2DM, Type 2 diabetes mellitus. **Note:** [a], Systematic reviews and meta-analyses of RCTs were identified and flagged. Bibliographies of these systematic reviews will be screened to check if literature searches have missed any potentially relevant studies.		

Two reviewers independently inspected each reference (title and abstract) identified by the literature searches and applied basic study selection criteria based on the eligibility criteria in Table 10 (primary screening). Citations meeting basic study selection criteria (or in cases of disagreement between the two reviewers) were obtained in full and independently assessed against the full eligibility criteria presented in Table 10 (secondary screening). In the event of disagreement between the two reviewers, a third reviewer independently assessed the paper and applicability of selection criteria was attained by consensus.

Where multiple publications were identified for the same clinical trial, all were included in the final list of articles meeting the eligibility criteria but clearly identified as primary and secondary sources for the same trial. All relevant data were extracted from the included full text of articles by one reviewer and quality checked against the original source by a second reviewer. Where more than one publication was identified describing a single trial, the data were compiled into a single entry in the data extraction table to avoid double counting of the patients. Each publication was referenced in the table to recognise that more than one publication may have contributed to the entry.

Search results

A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram showing the number of studies included and excluded at each stage of the review is presented in Figure 2.

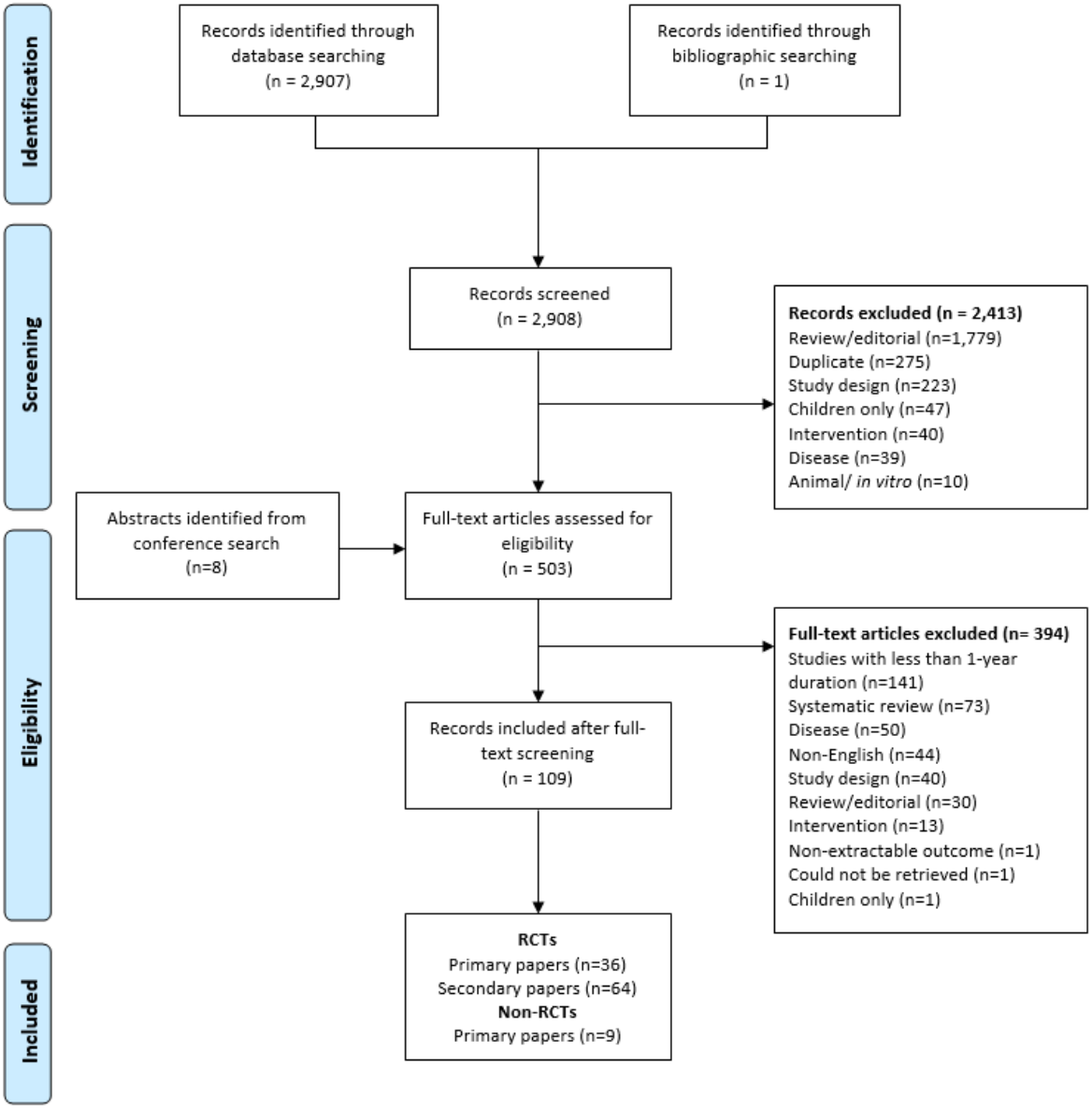
Database searches identified a total of 2,907 records. One additional study was identified through bibliographic searches of relevant systematic reviews. After primary screening of title and abstract, 2,413 records were excluded as they were not relevant to the research question. A total of 495 records were accessed in full. A further 8 conference abstracts were included for assessment at this stage resulting in a total of 503 records. After secondary screening, 394 records were excluded for

reasons such as study design or intervention not of interest resulting in a total of 109 records included in the review.

Of the 109 included records, 36 were primary publications of randomised controlled trials (RCTs), 64 were secondary publications of RCTs and 9 were non-RCT studies. Non-RCT studies were not considered further. Of the 36 included RCTs, 5 studies investigated treatment with NB32 (detailed in Table 11), while the remaining 31 studies investigated treatment with orlistat.

The orlistat studies have been used for comparative efficacy analyses and are therefore presented in Section 4.10.

Figure 2: PRISMA flow diagram of the clinical effectiveness literature search process (May 2016)



Key: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-analyses; RCT, randomised controlled trial.

4.2 List of relevant randomised controlled trials

A summary of the studies investigating NB that were identified through systematic review are presented in Table 11. Of note, two additional Phase II, dose-response studies were included in the regulatory file, but were excluded from the SLR due to a trial duration of less than 1 year. These data have not been used in the comparative

efficacy or cost-effectiveness assessments; however, further details of these studies are available on request.

Table 11: List of relevant RCTs

Trial name (NCT number)	Population	Intervention	Comparator	Primary study reference
COR-I (NCT00532779)	Adults with uncomplicated obesity or who were overweight with dyslipidaemia or hypertension	Naltrexone 32mg per day + bupropion 360mg per day (NB32) Naltrexone 16mg per day + bupropion 360mg per day (NB16)	Placebo	Greenway et al. 2010[16]
COR-BMOD (NCT00456521)	Adults with uncomplicated obesity or who were overweight with dyslipidaemia or hypertension	Naltrexone 32mg per day + bupropion 360mg per day (NB32) + BMOD	Placebo + BMOD	Wadden et al. 2011[18]
COR-II (NCT00567255)	Adults with uncomplicated obesity or who were overweight with dyslipidaemia or hypertension	Naltrexone 32mg per day + bupropion 360mg per day (NB32)	Placebo	Apovian et al. 2013[17]
COR-DM (NCT00474630)	Adults with T2DM and BMI ≥27 and ≤45kg/m^2	Naltrexone 32mg per day + bupropion 360mg per day (NB32)	Placebo	Hollander et al. 2013[19]
NB-CVOT (NCT01601704)	Adults with a BMI of 27–50 and who had characteristics associated with an increased risk of CV outcomes	Naltrexone 32mg per day + bupropion 360mg per day (NB32)	Placebo	Nissen et al. 2016[52]

Key: BMI, body mass index; BMOD, intensive behaviour modification; COR, Contrave obesity research; CV, cardiovascular; CVOT, cardiovascular outcomes trial; DM, diabetes mellitus; RCTs, randomised controlled trials; T2DM, Type 2 diabetes mellitus.

The clinical trial programme evaluating the efficacy of NB32 included four pivotal, Phase III studies, across which a total of 2,510 patients were randomised to NB32 treatment, and 1,448 patients were randomised to placebo. The omission of an

active pharmacological parameter was deemed acceptable by the EMA due to distinct tolerability profiles making blinding of medicine difficult[53]; this is discussed further in Section 4.13. However, in all trials, patients in both arms did receive some form of intervention: customary diet and behaviour modification in the case of COR-I, COR-II and COR-DM, and intensive behaviour modification in the case of COR-BMOD. These measures are thought to reflect the varying levels of non-pharmacological treatment approaches adopted in clinical practice, and thus provide direct trial data for the comparison of NB32 plus standard management versus standard management without NB32.

The NB-CVOT study was primarily designed to investigate the CV safety of NB32 in weight management (see Section 2.2), but also provides longer-term efficacy data. It should be acknowledged that these data were analysed despite the study being terminated earlier than originally planned (after the 50% interim analysis), after 25% interim data were made public in a US patent (and related Orexigen security filings) and by the EMA in the Mysimba EPAR. An additional, 78-week RCT (the IGNITE study) is also discussed in Section 4.7.[54] At the time of database searches, this study was not yet published and was therefore not identified or included in the SLR.

This submission focuses on data from the four pivotal RCTs: COR-I, COR-II, COR-BMOD and COR-DM, with only longer-term efficacy and safety data used to predict maintenance of pivotal trial outcomes presented from the NB-CVOT study and supported with data from the IGNITE study.

4.3 Summary of methodology of the relevant randomised controlled trials

A summary of the four pivotal RCTs discussed within this submission is presented in Table 12.

The efficacy, safety and tolerability of NB32 was evaluated in overweight and obese patients receiving customary diet and behaviour modification. This included a hypocaloric diet (500 kilocalorie [kcal] per day deficit based on the World Health Organization [WHO] algorithm for calculating resting metabolic rate) as well as instructions on increasing physical activity (COR-I and COR-II), or more intensive behaviour modification counselling (COR-BMOD). In addition, as recommended by

the current CHMP guideline on clinical evaluation of medicinal products used in weight control[55], a further study was conducted in obese/overweight patients with T2DM (COR-DM). All studies included a NB32 and placebo treatment arm. In the COR-I study, patients could also be randomised to a fixed oral dose of sustained-release (SR) 16mg per day naltrexone plus SR 360mg per day bupropion (NB16: [4mg naltrexone/90mg bupropion in each tablet, two tablets taken twice a day]).[16] In the COR-II study, to evaluate the efficacy and safety of a dose increase in participants with suboptimal response, NB32 participants with <5% weight loss at visits between Weeks 28 and 44 inclusive were re-randomised in a double-blind 1:1 ratio to continue receiving NB32 or escalate to NB48 (48mg/day naltrexone SR plus 360mg/day bupropion SR) for the remainder of the study.[17] Each trial included a 4-week dose escalation period, beginning with a quarter of the full dose, which increased weekly to Week 4, after which full dosing was maintained throughout 52 weeks of treatment. In the COR-I study, after 56 weeks of treatment, patients were re-randomised in a double blind 1:1 ratio to undergo tapered or sudden withdrawal of study drug.[56]

All four studies employed the FDA recommended co-primary endpoints of percent change in body weight from baseline and the proportion of patients who achieved ≥5% weight loss; this was assessed at Week 56 in all studies except COR-II, where primary analysis was conducted at Week 28 due to the potential to re-randomisation to a higher dose in non-responders after the Week 28 assessment.

Secondary endpoints included the proportion of patients who achieved ≥10% weight loss (recommended by the CHMP)[55], and various cardiometabolic parameters including:

- Waist circumference, a measurement that further characterises change in adiposity with improvements observed in reduced measurements[2, 5];
- Lipid levels, that are cardioprotective and may decrease the risk of coronary artery and cerebrovascular disease with improvements observed in reduced levels of LDL cholesterol and triglycerides and increased levels of HDL cholesterol[57];

- High-sensitivity C-reactive protein (hs-CRP), used to measure inflammation that may be associated with CV risk with improvements observed in reduced levels.[58]

Further secondary endpoints included fasting insulin levels, insulin resistance (quantified by the homeostasis model assessment of insulin resistance [HOMA-IR] assessment), and fasting blood glucose, which are often used to predict risk of diabetes[59, 60] in the 'pre-diabetic' populations of COR-I, COR-II and COR-BMOD (higher levels indicate increased risk); and HbA1c levels, which are used to monitor glycaemic control in the diabetic population of COR-DM with higher HbA1c levels indicating poorer control. HRQL was also assessed through the IWQOL-Lite, COE, FCI and IDS-SR questionnaires. In the COE questionnaire, item #19, which asks '*Generally, how difficult has it been to control your eating?'* was selected as the outcome of interest. Further details of these tools are presented in Appendix 6.

Of note, change in BMI was not a pre-defined endpoint. Although this is an adequate research tool, it is limited in the assessment of an individual, as it does not consider different body morphologies (e.g. muscle vs adipose) and may be skewed by very high muscle mass.[61] In addition, some population groups, such as people of Asian family origin and older people, have comorbidity risk factors that are of concern at different BMIs (lower for adults of an Asian family origin and higher for older people).[3] Therefore, alternative methods to measure body fatness, such as waist circumference, were utilised in the trials.

Table 12: Comparative summary of trial methodology for COR-I and COR-II

	COR-I	COR-II
Location	Patients were treated at 34 study sites in the US	Patients were treated at 36 study sites in the US
Trial design	A Phase III, multicentre, randomised, double-blind, placebo-controlled, 56-week study. Patients were randomised in a 1:1:1 ratio through a computer-generated, web-based system. Randomisation was stratified by study centre. Included a sub-study in which patients underwent body composition analysis and visceral fat measurement at baseline and after approximately 52 weeks of therapy	A Phase III, randomised, parallel-arm, double-blind, placebo-controlled, 56-week study. Patients were randomised in a 2:1 ratio through an interactive voice response system. Randomisation was stratified by study site. Included a sub-study in which blood pressure was measured over a 24-hour period at baseline, and after approximately 24 and 52 weeks of therapy.
Key eligibility criteria for patients	Patients aged 18–65 years; BMI 30–45kg/m^2 and uncomplicated obesity OR BMI 27–45kg/m^2 and controlled hypertension and/or dyslipidaemia were included. Patients were excluded if they met any of the following criteria: Type 1 or 2 diabetes; significant vascular, hepatic or renal disease; weight change of >4kg within 3 months prior to randomisation; history of seizures or serious psychiatric illness; obesity of known endocrine origin; history of malignancy within previous 5 years; bipolar disorder; history of drug or alcohol abuse or dependence within 1 year prior to study initiation; received excluded concomitant medication; history of surgical or device intervention for obesity; history of treatment with, hypersensitivity or intolerance to bupropion or naltrexone; initiation of discontinuation of tobacco products within 3 months prior to randomisation; females who were pregnant or breast-feeding or planning to become pregnant during the study period.	Patients aged 18–65 years old; BMI 30–45kg/m^2 and uncomplicated obesity OR BMI 27–45kg/m^2 and controlled hypertension and/or dyslipidaemia were included. Patients were excluded if they met any of the following criteria: Type 1 or 2 diabetes; significant vascular, hepatic or renal disease; weight change of >4kg within 3 months prior to randomisation; history of seizures or serious psychiatric illness; obesity of known endocrine origin; history of malignancy within previous 5 years; bipolar disorder; history of drug or alcohol abuse or dependence within 1 year prior to study initiation; received excluded concomitant medication; history of surgical or device intervention for obesity; history of treatment with, hypersensitivity or intolerance to bupropion or naltrexone; initiation of discontinuation of tobacco products within 3 months prior to randomisation; females who were pregnant or breast-feeding or planning to become pregnant during the study period.

	COR-I	COR-II
Trial drugs	NB32 (n=583): Naltrexone 32mg per day + bupropion 360mg per day; two tablets to be taken twice daily (each tablet contains 8mg naltrexone hydrochloride and 90mg bupropion hydrochloride). NB16 (n=578): Naltrexone 16mg per day + bupropion 360mg per day; two tablets to be taken twice daily (each tablet contains 4mg naltrexone hydrochloride and 90mg bupropion hydrochloride) Placebo (n=581): two tablets to be taken twice daily At baseline, 12, 24, 36 and 48 weeks, patients received instructions to follow a hypocaloric diet (500 kcal/day deficit) and increase physical activity, and behaviour modification advice. After a 4-week dose escalation period, treatment was continued for 52 weeks. Patients were free to discontinue their participation at any time.	NB32 (n=1,001): Naltrexone 32mg per day + bupropion 360mg per day; two tablets to be taken twice daily (each tablet contains 8mg naltrexone hydrochloride and 90mg bupropion hydrochloride). Placebo (n=495): two tablets to be taken twice daily NB32 patients with <5% weight loss at visits between weeks 28 and 44 inclusive were re-randomised (double-blind, 1:1 ratio) to continue receiving NB32 or escalate to NB48. From Week 29: NB48 (n=123): Naltrexone 48mg per day + bupropion 360mg per day NB32 (n=128): Naltrexone 32mg per day + bupropion 360mg per day Placebo (n=495) At baseline, 12, 24, 36 and 48 weeks, patients received instructions to follow a hypocaloric diet (500 kcal/day deficit) and increase physical activity, and behaviour modification advice. After a 4-week dose escalation period, treatment was continued for 52 weeks. Patients were free to discontinue their participation at any time.
Permitted and disallowed concomitant medication	Psychotropic agents with the exception of low-dose benzodiazepine or hypnotic agents for the treatment of insomnia; anorectic or weight loss agents; alpha-adrenergic blockers and clonidine; dopamine agonists; Coumadin; theophylline; cimetidine; oral corticosteroids; cholestyramine or cholestypol; Depo-provera®; smoking cessation agents and use of opioid or opioid-like analgesics were all prohibited.	Psychotropic agents with the exception of low-dose benzodiazepine or hypnotic agents for the treatment of insomnia; anorectic or weight loss agents; alpha-adrenergic blockers and clonidine; dopamine agonists; Coumadin; theophylline; cimetidine; oral corticosteroids; cholestyramine or cholestypol; Depo-provera®; smoking cessation agents and use of opioid or opioid-like analgesics were all prohibited.

	COR-I	COR-II
	Anti-hypertensive medications were allowed at study entry if the regimen had been stable for 6 weeks and the patient's systolic blood pressure was ≤140mmHg and diastolic blood pressure was ≤90mmHg. Medications for the treatment of dyslipidaemia were allowed at study entry if the regimen had been stable for 6 weeks and the patient's triglycerides level was <400mg/dL. Treatment of nausea and insomnia was permitted.	Anti-hypertensive medications were allowed at study entry if the regimen had been stable for 6 weeks and the patient's systolic blood pressure was ≤140mmHg and diastolic blood pressure was ≤90mmHg. Medications for the treatment of dyslipidaemia were allowed at study entry if the regimen had been stable for 6 weeks and the patient's triglycerides level was <400mg/dL. Treatment of nausea and insomnia was permitted.
Primary outcomes	Percentage of change in total body weight and proportion of patients with ≥5% decrease in total body weight at Week 56.	Percentage of change in total body weight and proportion of patients with ≥5% decrease in total body weight at Week 28.
Secondary outcomes	Change in the following variables from baseline to Week 56: Proportion of patients with ≥10% decrease in total body weight; waist circumference; fasting HDL; fasting triglycerides; IWQOL-Lite total score; hs-CRP; fasting insulin; fasting blood glucose; HOMA-IR; 21-item COE questionnaire item no. 19; fasting LDL; systolic blood pressure; diastolic blood pressure; IDS-SR total score[a]; FCI sweets subscale and carbohydrates/starches subscale scores; safety including AEs, TEAEs, SAEs; laboratory data; vital signs including blood pressure and pulse rate.	Change in the following variables from baseline to Week 56: Percent change in total body weight (using weighted LOCF analysis); proportion of patients with ≥5% decrease in total body weight (using weighted LOCF analysis). Change in the following variables from baseline to Week 28: Proportion of patients with ≥10% decrease in total body weight; waist circumference; fasting HDL; fasting triglycerides; IWQOL-Lite total score; hs-CRP; fasting insulin; fasting blood glucose; HOMA-IR; 21-item COE questionnaire item no. 19; fasting LDL; systolic blood pressure; diastolic blood pressure; IDS-SR total score[a]; FCI sweets subscale and carbohydrates/starches subscale scores; safety including AEs, TEAEs, SAEs; laboratory data; vital signs including blood pressure and pulse rate. Additional analyses were conducted on the co-primary endpoints and a selected number of secondary variables at Week 56, assessing the pooling of all NB32- and NB48-treated subjects compared to placebo. These analyses

	COR-I	COR-II
		were not part of the closed testing procedure. The secondary variables included waist circumference, IDS-SR total score, HDL cholesterol, total cholesterol, triglycerides, pulse rate, systolic and diastolic blood pressure values, and FCI sweets and carbohydrates subscales scores.
Tertiary/ Exploratory outcomes	Change in the following variables from baseline to Week 56: Change in total body weight; proportion of patients with ≥10% decrease in total body weight in the completers analysis set and ITT analysis set; pulse rate; IWQOL-Lite subscale scores including physical function, self-esteem, sexual life, public distress and work; FCI questionnaire total score and subscale score; 21-item COE questionnaire subscale score; depressive symptoms measured by total score on the IDS-SR scale[a]; time to response (≥5% weight loss from baseline) using KM estimates. Change in the following variables from baseline to Week 28: Percentage change in total body weight; change in total body weight; proportion of patients with ≥10% decrease in total body weight in the mITT; selected obesity-associated CV risk factors including serum triglycerides, fasting insulin, fasting blood glucose and systolic and diastolic blood pressures; other obesity-associated CV risk factors including waist circumference, pulse rate, HDL and LDL cholesterol, HOMA-IR, hs-CRP; IWQOL-Lite total score and subscale scores for physical function, self-esteem, sexual life, public distress and work; FCI questionnaire total score and subscale score; 21-item COE questionnaire subscale score; depressive symptoms	Percent change in change from baseline in total body weight by visit; proportion of patients with ≥5% decrease in total body weight by visit; Change in the following variables from baseline to Week 56: Proportion of patients with ≥10% decrease in total body weight for the completers analysis set, ITT analysis set and the mITT; selected obesity-related CV risk factors including serum triglycerides, fasting insulin, fasting blood glucose, systolic and diastolic blood pressure; other obesity-associated CV risk factors including waist circumference, pulse rate, HDL and LDL cholesterol, hs-CRP and HOMA-IR; IWQOL-Lite total score; IWQOL-Lite subscale scores for physical function, self-esteem, sexual life, public distress and work, by visit; change in SF-36 MCS and PCS and individual item scores; FCI questionnaire total score and subscale scores by visit; 21-item COE by visit; depressive symptoms measured by the total score on the IDS-SR scale[a]; time to response (≥5% weight loss from baseline) using KM estimates. Change in the following variables from baseline to Week 28: Proportion of patients with ≥10% decrease in total body weight for the completers analysis set and the ITT analysis set; pulse rate; IWQOL-Lite subscale scores for physical function, self-esteem, sexual life, public distress and work, by visit; change in SF-36 MCS and PCS and individual

	COR-I	COR-II
	measured by total score on the IDS-SR scale[a]; total cholesterol.	item scores; FCI questionnaire total score and subscale scores by visit; 21-item COE by visit; total cholesterol.
Pre-planned subgroups	For the co-primary efficacy variable, subgroup analyses were conducted within selected subpopulations defined by factors including study centre, sex, race, age, age group, BMI category, and tobacco use.	For the co-primary efficacy variable, subgroup analyses were conducted within selected subpopulations defined by factors including study centre, sex, race, age, age group, BMI category, presence of hypertension and/or dyslipidaemia and tobacco use.

Key: BMI, body mass index; COE, Control of Eating questionnaire; CV, cardiovascular; FCI, Food Craving Inventory; HDL, high-density lipoprotein; HOMA-IR, homeostasis model assessment of insulin resistance; hs-CRP, high-sensitivity C reactive protein; IDS-SR, Inventory of Depressive Symptoms – Subject related; ITT, intent-to-treat; IWQOL-Lite, Impact of Weight on Quality of Life-Lite version; LOCF, last observation carried forward; LDL, low-density lipoprotein; MCS, mental component score; mITT, modified intent-to-treat; PCS, physical component score.
Notes: [a], The IDS-SR questionnaire was used as both an efficacy and safety variable.
Source: Apovian et al. 2013[17]; Greenway et al. 2010[16]; Orexigen, 2010[56]; Orexigen, 2010[62]

Table 13: Comparative summary of trial methodology for COR-BMOD and COR-DM

	COR-BMOD	COR-DM
Location	Patients were treated at nine study sites in the US.	Patients were treated at 53 study sites in the US.
Trial design	A Phase III, multicentre, randomised, double-blind, placebo-controlled, 56-week study. Patients were randomised in a 3:1 ratio (NB32:placebo) via a centralised automated voice response system. Randomisation was stratified by study centre.	A Phase III, randomised, double-blind, placebo-controlled, 56-week study. Patients were randomised in a 2:1 ratio (NB32:placebo) via a computer-generated randomisation schedule. Randomisation was stratified by baseline HbA1c (≤8 or >8%; ≤64 or >64mmol/mol) and sulfonylurea use.
Key eligibility criteria for patients	Patients aged 18–65 years; BMI 30–45kg/m^2 and uncomplicated obesity OR BMI 27–45kg/m^2 and controlled hypertension and/or dyslipidaemia were included. Patients were excluded if they met any of the following criteria:	Patients were included if they met the following criteria: Smoking or non-smoking men and women with T2DM; aged between 18–70 years; BMI ≥27 and ≤45kg/m^2; HbA1c between 7% and 10% and fasting blood glucose <270mg/dL; not taking a diabetes medication or were on

	COR-BMOD	COR-DM
	Type 1 or 2 diabetes; significant vascular, hepatic or renal disease; weight change of >4kg within 3 months prior to randomisation; history of seizures or serious psychiatric illness; obesity of known endocrine origin; history of malignancy; bipolar disorder; history of drug or alcohol abuse or dependence within 1 year prior to study initiation; received excluded concomitant medication; history of surgical or device intervention for obesity; history of treatment with, hypersensitivity or intolerance to bupropion or naltrexone; use of tobacco products within 6 months prior to screening; females who were pregnant or breast-feeding or planning to become pregnant during the study period.	stable doses or oral antidiabetes drugs for ≥3 months prior to randomisation; systolic and diastolic blood pressure of <145 and <95 mmHg, respectively. Patients were excluded if they met the following criteria: Type 1 diabetes; obesity of known endocrine origin other than diabetes mellitus; diabetes mellitus secondary to pancreatitis or pancreatectomy; significant vascular, hepatic or renal disease; history of malignancy within 5 years prior to screening; loss or gain of more than 5.0kg within 3 months prior to screening; severe micro- or macrovascular complications of diabetes; serious psychiatric illness; bipolar disorder; history of drug or alcohol abuse or dependence within 1 year prior to screening; history of surgical or device intervention for obesity; history of seizures of any aetiology; treatment with bupropion or naltrexone within 12 months prior to screening; history of hypersensitivity or intolerance to bupropion or naltrexone; change in smoking status in the previous 3 months; females who were pregnant or breast-feeding or planning to become pregnant during the study period.
Trial drugs	NB32 + BMOD (n=591): Naltrexone 32mg per day + bupropion 360mg; two tablets to be taken twice daily (each tablet contains 8mg naltrexone hydrochloride and 90mg bupropion hydrochloride). BMOD consisted of group meetings lasting 90 minutes weekly for the first 16 weeks, every other week for the next 12 weeks and monthly thereafter. They included instructions to consume a balanced deficit diet and to increase to 180 min/week of planned, moderately vigorous, physical activity. Placebo + BMOD (n=202): two placebo pills to be taken twice daily + BMOD as described above	NB32 (n=335): Naltrexone 32mg per day + bupropion 360mg per day; two tablets to be taken twice daily (each tablet contains 8mg naltrexone hydrochloride and 90mg bupropion hydrochloride). Placebo (n=170): two placebo pills to be taken twice daily At baseline and Weeks 4, 16, 28, and 40, all participants were instructed by study site personnel to follow a hypocaloric diet (500 kcal deficit/day, based on the World Health Organization algorithm for calculating resting metabolic rate). Participants received dietary counselling and the “Exchange Lists for Weight Management”

	COR-BMOD	COR-DM
	After a 4-week dose escalation period, treatment was continued for 52 weeks. Patients were free to discontinue their participation at any time.	booklets in accordance with the American Diabetes Association and American Dietetic Association guidelines. Participants also received advice on behaviour modification, including written instructions, to increase physical activity (to walking for at least 30 min most days of the week). After a 4-week dose escalation period, treatment was continued for 52 weeks. Patients were free to discontinue their participation at any time.
Permitted and disallowed concomitant medication	Psychotropic agents with the exception of low-dose benzodiazepine or hypnotic agents for the treatment of insomnia; anorectic or weight loss agents; alpha-adrenergic blockers and clonidine; Coumadin; theophylline; cimetidine; oral corticosteroids; topiramate; Depo-provera®; smoking cessation agents and use of opioid or opioid-like analgesics were all prohibited. Anti-hypertensive medications were allowed at study entry if the regimen had been stable for 8 weeks and the patient's systolic blood pressure was ≤140mmHg and diastolic blood pressure was ≤90mmHg. Medications for the treatment of dyslipidaemia were allowed at study entry if the regimen had been stable for 8 weeks and the patient's triglycerides level was <400mg/dL and LDL <190 /dL. Treatment of nausea and insomnia was permitted.	Psychotropic agents with the exception of low-dose benzodiazepine or hypnotic agents for the treatment of insomnia; anorectic or weight loss agents; alpha-adrenergic blockers, beta-blockers, dopamine agonists and clonidine; Coumadin; theophylline; cimetidine; oral corticosteroids; cholestypol or cholestyramine; Depo-provera®; smoking cessation agents; use of opioid or opioid-like analgesics were all prohibited. Anti-hypertensive medications were allowed at study entry if the regimen had been stable for 4 weeks. Medications for the treatment of dyslipidaemia were allowed at study entry if the regimen had been stable for 4 weeks and the patient's triglycerides level was <400mg/dL. Treatment of nausea and insomnia, as well as antidiabetic agents was permitted.
Primary outcomes	Percentage of change in total body weight and proportion of patients with ≥5% decrease in total body weight at Week 56.	Percentage of change in total body weight and proportion of patients with ≥5% decrease in total body weight at Week 56.
Secondary outcomes	Change in the following variables from baseline to Week 56:	Change in the following variables from baseline to Week 56:

	COR-BMOD	COR-DM
	Proportion of patients with ≥10% decrease in total body weight; waist circumference; fasting HDL; fasting triglycerides; IWQOL-Lite total score; hs-CRP; fasting insulin; fasting blood glucose; HOMA-IR; 21-item COE questionnaire; fasting LDL; systolic blood pressure; diastolic blood pressure; IDS-SR total score[a]; FCI sweets subscale and carbohydrates/starches subscale scores; safety including AEs, TEAEs, SAEs; laboratory data; vital signs including blood pressure and pulse rate.	HbA1c; fasting triglycerides; HDL cholesterol; blood glucose; waist circumference; proportion of patients with ≥10% decrease in total body weight; HbA1c <7%; percent of patients requiring rescue medications for diabetes; percent of patients requiring change in dose of oral anti-diabetes medication; HOMA-IR; fasting insulin; HbA1c <6.5%; IWQOL-Lite total score; hs-CRP; patients discontinuing due to poor glycaemic control; COE questionnaire item no. 19; fasting LDL cholesterol; systolic blood pressure; diastolic blood pressure; IDS-SR total score[a]; FCI sweets subscale and carbohydrates/starches subscale; safety including AEs, TEAEs, SAEs; laboratory data; vital signs including blood pressure and pulse rate.
Tertiary/ Exploratory outcomes	Change in the following variables from baseline to Week 56: IWQOL-Lite subscale scores for physical function, self-esteem, sexual life, public distress and work; FCI questionnaire total score and subscale scores by visit; time to response (≥5% weight loss from baseline) using KM estimates. Change in the following variables from baseline to Week 28: Percent and change in total body weight; proportion of patients with ≥5% decrease in total body weight; proportion of patients with ≥10% decrease in total body weight; selected obesity-related CV risk factors including serum triglycerides, fasting insulin, fasting blood glucose, systolic and diastolic blood pressure; other obesity-associated CV risk factors including waist circumference, pulse rate, HDL and LDL cholesterol, hs-CRP and HOMA-IR; IWQOL-Lite total score; IWQOL-Lite subscale scores	Change in the following variables from baseline to Week 56: Change in total body weight; proportion of patients with ≥10% decrease in total body weight for the completer and ITT analysis set; pulse rate; IWQOL-Lite subscale scores for physical function, self-esteem, sexual life, public distress and work; FCI questionnaire total score and subscale scores; 21-item COE; time to response (≥5% weight loss from baseline) using KM estimates. Change in the following variables from baseline to Week 28: Percent and change in total body weight; proportion of patients with ≥5% decrease in total body weight; change in HbA1c; proportion of patients with ≥10% decrease in total body weight; selected obesity-related CV risk factors including serum triglycerides, fasting insulin, fasting blood glucose, systolic and diastolic blood pressure; other obesity-associated CV risk factors including LDL and HDL

	COR-BMOD	COR-DM
	for physical function, self-esteem, sexual life, public distress and work; FCI questionnaire total score and subscale scores; 21-item COE; IDS-SR total score[a].	cholesterol; HOMA-IR, hs-CRP and waist circumference; pulse rate; IWQOL-Lite total score; IWQOL-Lite subscale scores for physical function, self-esteem, sexual life, public distress and work; FCI questionnaire total score and subscale scores; 21-item COE; IDS-SR total score[a]; total cholesterol.
Pre-planned subgroups	Subgroup analyses were performed to assess the effects of combination treatment within selected special populations defined by factors such as study centre, race, sex, age, and BMI categorisation.	For the co-primary efficacy variable, subgroup analyses were conducted within selected subpopulations defined by factors such as study centre, sex, race, age, age group, BMI category, presence of hypertension and/or dyslipidaemia; tobacco use; HbA1c strata, and sulfonylurea pharmacotherapy.

Key: BMI, body mass index; COE, Control of Eating questionnaire; FCI, Food Craving Inventory; HDL, high-density lipoprotein; HOMA-IR, homeostasis model assessment of insulin resistance; hs-CRP, high-sensitivity C reactive protein; IDS-SR, Inventory of Depressive Symptoms – Subject related; IVRS, interactive voice response system; IWQOL-Lite, Impact of Weight on Quality of Life-Lite version; KM, Kaplan–Meier; LDL, low-density lipoprotein; TEAE, treatment-emergent event; T2DM, Type 2 diabetes mellitus.
Notes: [a], IDS-SR questionnaire was used as both an efficacy and safety variable.
Source: Wadden et al. 2011[18]; Hollander et al. 2013[19]; Orexigen, 2010[63]; Orexigen, 2009[64]

The NB-CVOT Study

The NB-CVOT study was a Phase IIIb, multicentre, randomised, double-blind, placebo-controlled trial to assess the occurrence of MACE in overweight or obese patients.[52] Results within this submission include the 50% interim data and data accumulated after the February 2015 database lock, which report 64% of planned events.

Following a double-blind lead-in period, patients were randomised in a 1:1 ratio through an interactive voice recognition system to receive treatment with NB32 or placebo; the study included a 4-week dose escalation period followed by a maintenance period. Patients were also encouraged to participate in an internet-based weight management program as well as having access to a personal weight loss coach and a low-fat, low-calorie meal plan. At 16 weeks, if patients did not lose ≥2% of their initial body weight or experienced a sustained (at ≥2 visits) increase in blood pressure (systolic or diastolic) of 10mmHg or greater they were discontinued.[52] Further discussion is provided in Section 4.13.

Patients were eligible for inclusion in the study if they were aged 45 (men) or 50 (women) years or older, had a BMI 27–50kg/m^2 and a waist circumference of 88cm (women) or 102cm (men) or more. Enrolment was restricted to patients with characteristics associated with an increased risk of adverse CV outcomes. Patients were excluded for a myocardial infarction within 3 months prior to screening, severe angina pectoris, New York Heart Association class 3 or 4 heart failure, or history of stroke, or blood pressure of 145/95mmHg or higher. Patients were also excluded for unstable weight within 3 months prior to screening (weight gain or loss of >3%), planned bariatric or cardiac surgery, or percutaneous coronary intervention.

The prespecified primary outcome measure was time from treatment randomisation to the first confirmed occurrence of a MACE, defined as CV death, nonfatal stroke, or nonfatal myocardial infarction. Secondary outcomes included time to first MACE, stroke or myocardial infarction. A summary of participant flow is presented in Section 4.5, and outcome data are presented in Sections 4.7 and 4.12. Additional discussion of the statistical analysis, participant flow and quality assessment of the NB-CVOT study is presented in Appendix 3

The IGNITE Study

IGNITE was a Phase IIIb, randomised, open-label, controlled study in which patients received NB32 plus comprehensive lifestyle intervention (CLI) or usual care (standard diet and exercise advice) for 26 weeks.[54] NB32 + CLI patients not achieving 5% weight loss at Week 16 were discontinued, as indicated by product labelling. After Week 26, usual care patients began NB32 + CLI. Assessments continued through Week 78. The primary endpoint was percent change in weight from baseline to Week 26 in the per protocol (PP) population. Other endpoints included percentage of patients achieving ≥5%, ≥10% and ≥15% weight loss, percent change in weight at Week 78, AEs necessitating study discontinuation.

4.4 Statistical analysis and definition of study groups in the relevant randomised controlled trials

The hypothesis and associated statistical analysis methods adopted in the four pivotal RCTs are presented in Table 14.

Across all four pivotal trials, the primary analysis population was the modified intent-to-treat (mITT) population, defined as all randomised patients with a post-baseline body weight measurement obtained while the patient remained on study medication. Missing data was imputed using the LOCF method for primary analysis. Additional analysis populations included the intent-to-treat (ITT) population, defined as all randomised patients with a post-baseline body weight measurement; the per-protocol (PP) population, defined as all randomised patients who received at least 28 weeks of study treatment, were compliant with study medication [≥70% compliant], had a baseline measurement and had at least one post-baseline body weight measurement while on study drug; and the completers set. Additional methods of data imputation included repeated measures mixed effects, weight regain imputation and BOCF.

In all four pivotal trials, safety analyses were conducted on the safety analysis set. In the double-blind treatment phase, this included all randomised patients who were administered at least one tablet of study treatment and had at least one investigator contact/assessment at any time after the start of study treatment, regardless of whether they discontinue the study. In the drug discontinuation phase, the safety analysis set included all patients who were administered at least one tablet of study treatment and completed the drug discontinuation phase.[56, 62-64]

Table 14: Summary of statistical analyses in the RCTs

	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
COR-I	NB32 will result in greater weight loss than either treatment alone in adults with uncomplicated obesity or who are overweight with hypertension or dyslipidaemia	General linear models (ANCOVA) including terms for treatment and study centre and baseline values as covariates were used to analyse continuous endpoints. Categorical endpoints were analysed with a logistic regression model that included treatment and study centre as main effects and baseline bodyweight as a covariate. To maintain the family-wise type I error rate at 5%, secondary endpoints were analysed in a pre-determined sequence for each experimental group versus placebo. Formal testing was undertaken in a step-down manner until any endpoint failed to reach $p<0.05$, after which nominal p values are reported, and findings are deemed exploratory. To reduce skewness to a minimum, values for triglycerides, high-sensitivity C-reactive protein, insulin, and HOMA-IR were $\log_{10}$ transformed before running ANCOVA models. The percentage change from baseline was calculated by back-transforming the least squares geometric mean minus one.	The total sample size to be randomised was approximately 1,650 subjects with a 1:1:1 randomisation allocation between combination treatment and placebo groups. This sample size provided 99% power to detect a statistically significant difference between placebo and the combination treatment arms for the co-primary efficacy endpoints. The power calculation was made assuming the mean weight loss from baseline to the Week 56 visit would be approximately 1% for subjects randomised to placebo and ≥ 6% for subjects randomised to either combination treatment arm.	Primary analysis at 56 weeks was done with the LOCF on study drug. Additional methods for imputation included repeated measures mixed effects model, BOCF, and weight regain imputation., Patients without time-to-event analyses were considered right-censored.
COR-II	NB32 will result in greater weight loss than either treatment alone in adults with	General linear models (ANCOVA) including terms for treatment and study centre, with baseline values as covariates, were used to analyse the co-primary and	To obtain the targeted number of participant-exposures at 1 year, it was estimated that 1,000 participants would need to be	Missing data were imputed by carrying forward the last observation on

	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
	uncomplicated obesity or who are overweight with hypertension or dyslipidaemia	continuous secondary endpoints. Categorical endpoints were analysed using a logistic regression model including treatment and study centre as main effects and baseline values as covariates. To maintain the family-wise type I error rate at 5%, secondary endpoints were analysed in a predetermined sequence only after both co-primary endpoints achieved statistical significance. Formal testing was conducted in a step-down manner until any endpoint failed to reach $p<0.05$, after which the nominal p-values are reported and findings are considered exploratory. To control for skewness, analyses for triglycerides, hs-CRP, insulin, and HOMA-IR were log_{10} transformed prior to running the ANCOVA models. The percent change from baseline was calculated by back-transforming the LS geometric mean minus one. All statistical analyses were performed using Windows SAS version 9.1. Continuous endpoints are provided as LS mean ±SE unless otherwise indicated.	randomised to NB32, with an assumed 40% attrition rate, with a 99%, 81%, and 70% chance that >1 AE would be observed at a true frequency of 1/100, 1/250, and 1/500, respectively. It was estimated that 1,500 randomised participants (2:1 ratio) would provide 99% power to detect a statistically significant difference in mean percent weight loss of >5%, and a 14% difference in the proportion of participants with >5% weight loss between NB32 and placebo. Power estimates were determined using a two-sample t-test for mean percent weight loss and a two-sample continuity-corrected chi-square test for the proportion of participants with >5% weight loss using a two-sided significance level of 5%.	study drug (LOCF analysis) for primary analysis. Additional methods for imputation included repeated measures mixed effects model, BOCF, and weight regain imputation Patients without time-to-event analyses were considered right-censored.
COR-BMOD	NB32 will result in greater weight loss than either treatment alone in adults with uncomplicated obesity or who are overweight	Unless otherwise specified, when an ANCOVA model was used to analyse a continuous efficacy variable, the model contained treatment and study centre as main effects, and baseline values as covariates. Type III sums of squares for the LS means were used for the statistical	A total of 800 participants was determined to provide 99% power to detect a 5%-point difference between groups in percent change in initial weight, assuming the percentage change in weight for placebo + BMOD was ~5% (e.g. 5 vs	Missing data were imputed by carrying forward the last observation on study drug (LOCF analysis) for primary analysis.

	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
	with hypertension or dyslipidaemia	comparison. When a logistic regression model was used to analyse a categorical efficacy variable, the model contained treatment and study centre as main effects and baseline body weight value as a covariate. Categorical safety variables were analysed using Fisher's exact test, and continuous safety variables were analysed using an ANCOVA model with treatment and study centre as main effects, and the appropriate baseline values as covariates.	10%) on the first co-primary endpoint, and ~90% power to detect a 14%-point difference for the second co-primary endpoint, assuming the proportion of subjects achieving >5% weight loss was 50% in the placebo + BMOD group (e.g. 50 vs 64%).	Additional methods for imputation included repeated measures mixed effects model, BOCF, and weight regain imputation
COR-DM	The null hypotheses stated there were no differences between the treatment groups in the percent change in total body weight or the proportion of subjects with ≥5% decrease in total body weight from baseline to endpoint (Week 56)	General linear models (ANCOVA) including terms for treatment, HbA1c strata ≤8 or >8%, pharmacotherapy with or without sulfonylurea, and baseline values as covariates were used to analyse continuous endpoints. Categorical endpoints were analysed with a logistic regression model using the same covariates as the continuous endpoints. To minimise skewness, values for triglycerides, hs-CRP, insulin, and HOMA-IR were $\log_{10}$ transformed before running the general linear models. The LS percent change from baseline was calculated by back-transforming the LS mean in $\log_{10}$ scale. To control for multiple comparisons, secondary endpoints were analysed in a predetermined sequence. Testing	To obtain the targeted number of participant exposures at 1 year, the investigators estimated that 350 participants would need to be randomised to NB, with an assumed 33% attrition rate. It was estimated that 525 participants randomised 2:1 (~350 to NB and ~175 to placebo) would provide ~99% power to detect a difference in mean weight loss of ≥5% between NB and placebo (assuming an SD of 5%, comparison between groups using a two-sample t test and two-sided significance level of 0.05).	Missing data were imputed by carrying forward the last observation on study drug (LOCF analysis) for primary analysis. Additional methods for imputation included repeated measures mixed effects model, BOCF, and weight regain imputation.

	Hypothesis objective	Statistical analysis	Sample size, power calculation	Data management, patient withdrawals
		proceeded in a sequential step-down manner until any endpoint failed to reach p<0.05, after which nominal P values are reported and findings deemed exploratory. Continuous data are presented as LS mean ± SE unless otherwise indicated. All statistical analyses were performed using Windows SAS, version 9.1 (SAS Institute, Cary, NC).		

Key: ANCOVA, analysis of covariance; BMOD, behaviour modification; HOMA-IR, homeostasis model assessment – insulin resistance; hs-CRP, high-sensitivity C reactive protein; ITT, intent-to-treat; LOCF, last observation carried forward; LS, least squares; NB, naltrexone plus bupropion; SD, standard deviation; SE, standard error.
Source: Apovian et al. 2013[17]; Greenway et al. 2010[16]; Hollander et al. 2013[19]; Wadden et al. 2011[18]; Orexigen, 2010[62]; Orexigen, 2009[64]; Orexigen, 2010[56]; Orexigen, 2010[63]

4.5 Participant flow in the relevant randomised controlled trials

Participant flow

COR-I study

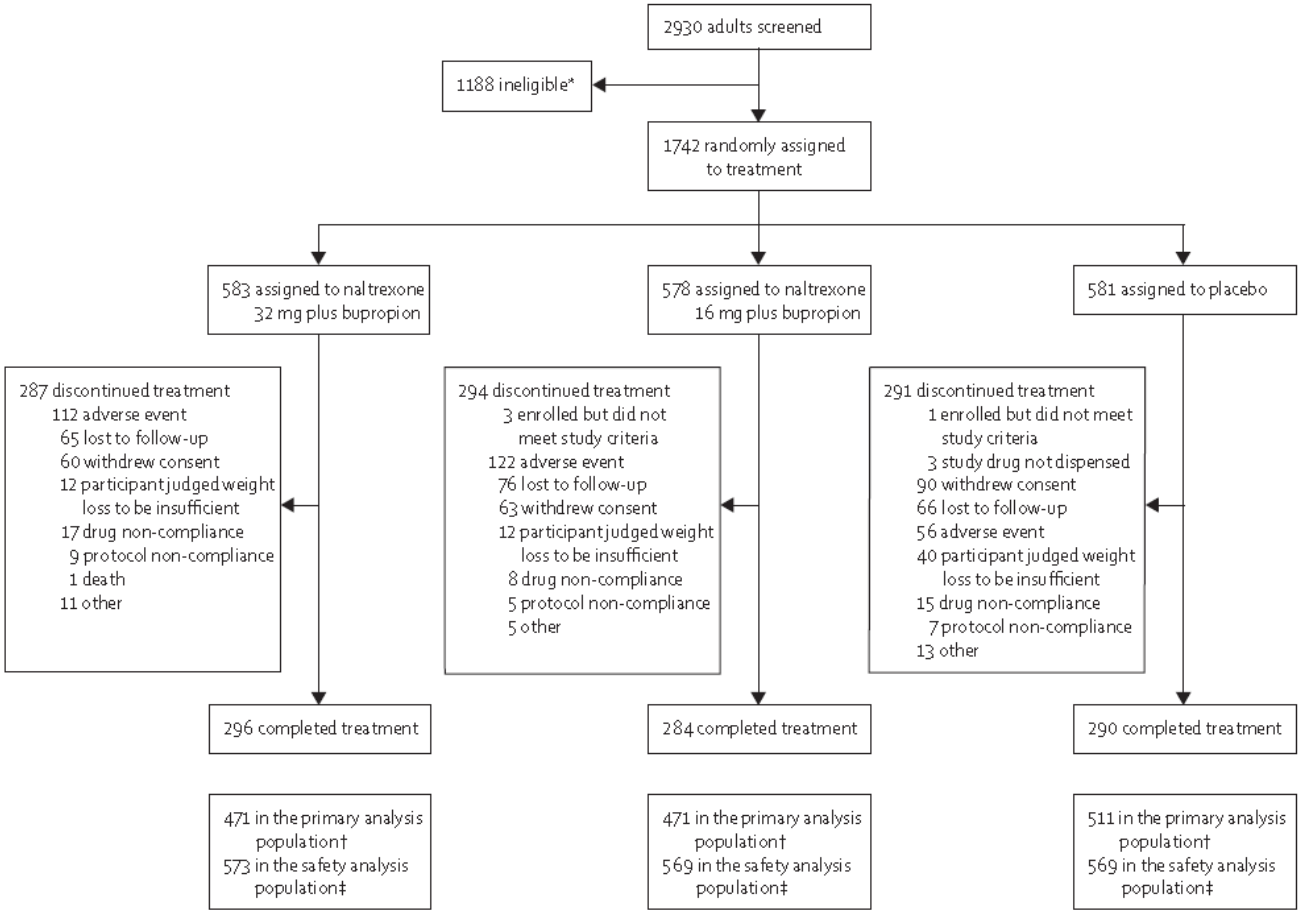
A total of 1,742 patients were randomised to treatment with NB32, NB16 or placebo. Of the 583 patients randomised to NB32, 471 (80.8%) qualified for inclusion in the mITT population; 578 patients were randomised to NB16, of which 471 (81.5%) qualified for inclusion in the mITT population; and 581 patients were randomised to placebo, of which 511 (88.0%) qualified for inclusion in the mITT population.[56]

Of all randomised patients, a total of 870 (50%) completed 56 weeks of treatment; 296 in the NB32 group, 284 in the NB16 group and 290 in the placebo group. Rates of discontinuation were similar across treatment groups. More patients in the NB groups discontinued because of AEs than patients in the placebo group ($p<0.0001$); discontinuation generally occurred early in the study (by Weeks 4 and 8). More patients in the placebo group discontinued because of insufficient weight loss ($p<0.0001$) and withdrawal of consent ($p=0.0126$) than patients receiving treatment with NB.[16] Rate of discontinuation was higher during the first 16 weeks of the study in both the placebo (180 of 291 patients who discontinued [61.9%]) and combination treatment groups (NB32: 204 of 287 [71.1%]; NB16: 218 of 294 [74.1%]).[16]

Participant flow for NB-301 is presented as a Consolidated Standards of Reporting Trials (CONSORT) diagram in Figure 3.

The mean duration of exposure to study drug was 34.2 weeks for the NB16 group, 35.5 weeks for the NB32 group and 36.1 weeks for the placebo group. The NB16, NB32 and placebo groups therefore represent a total of 373.7, 391.9 and 395.0 patient-years of exposure, respectively.[56] An exposure of ≤4 weeks was observed for 20.2% of patients in the NB16 group and 19.7% of patients in the NB32 group compared, with 10.4% of patients in the placebo group, consistent with a higher study drug discontinuation rate observed at Week 4 for the intervention groups over placebo.[56]

Figure 3: CONSORT diagram of participant flow in COR-I



Notes: *, Reasons for ineligibility of excluded adults are not available; †, The primary analysis population included all randomised patients with a baseline weight measurement and a post-baseline weight measurement while on study drug. Missing data were imputed by use of the last observation carried forward method; ‡, The safety analysis included all randomised patients who took one or more tablets of study drug and had at least one investigator contact or assessment any time after starting treatment.
Source: Greenway et al. 2010[16]

COR-II study

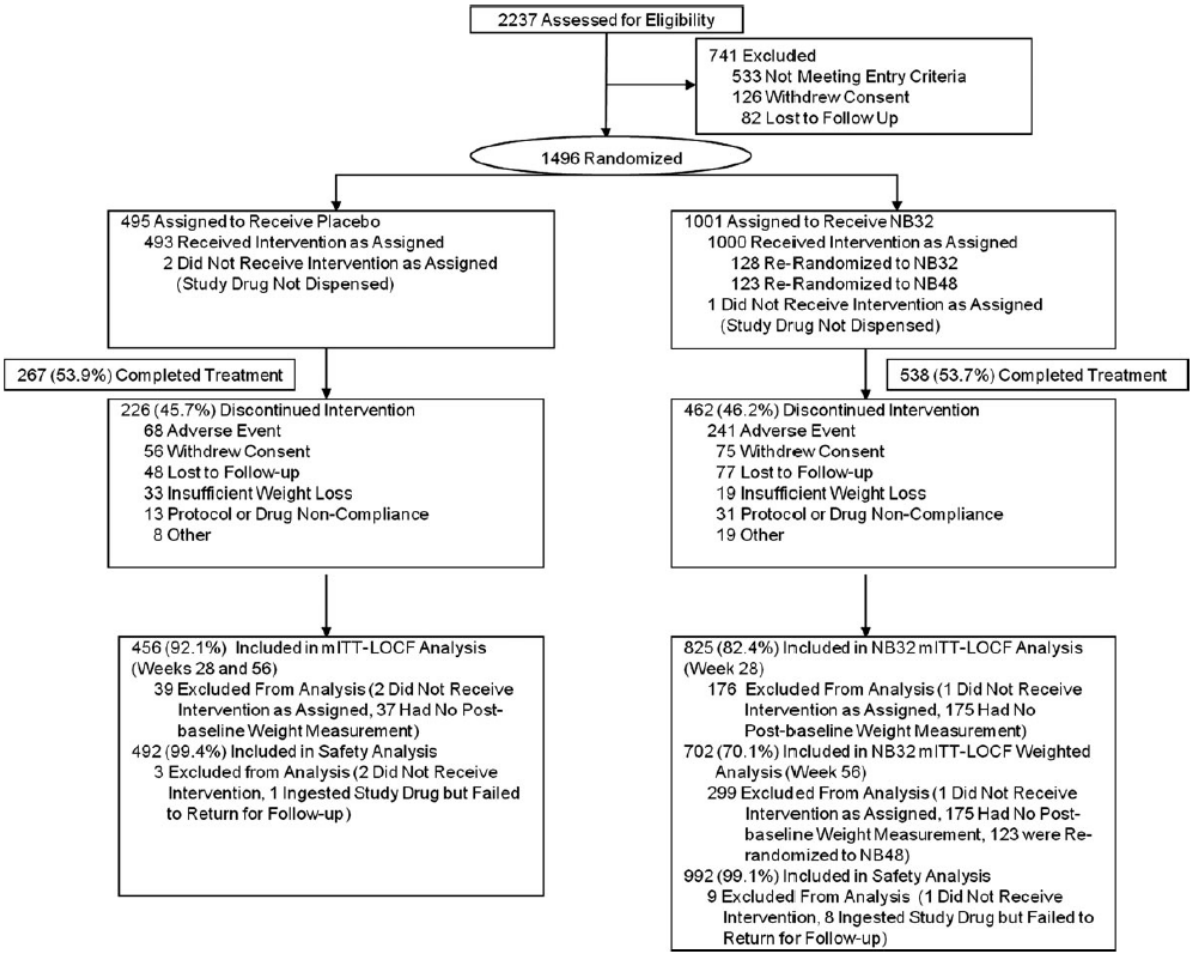
A total of 1,496 patients were randomised; 1,001 received NB32, and 495 received placebo. Of these patients, 825 (82.4%) patients in the NB32 group and 456 (92.1%) placebo treated patients were eligible for inclusion in the mITT population.[62]

Of the patients randomised to double-blind treatment, 54% of patients in each treatment group completed 56 weeks of treatment.[17] More NB32-treated patients discontinued because of an AE ($p<0.001$), whereas more placebo-treated patients discontinued because of insufficient weight loss ($p<0.001$) and withdrawal of consent ($p<0.05$). Discontinuations in both groups occurred most frequently during the first 8 weeks of the study, with more discontinuations, particularly because of AEs, occurring with NB32 treatment.[17]

A CONSORT flow diagram is presented in Figure 4.

The mean duration of exposure to study drug was 38.3 weeks for the placebo group, representing 357.9 patient years of exposure, and 36.4 weeks for the NB32 group, representing 690.1 patient years of exposure.[62] A total of 18% of patients in the NB32 group had ≤4 weeks of exposure compared with 7.9% of patients in the placebo group.

Figure 4: CONSORT diagram of participant flow in COR-II



Key: LOCF, last observation carried forward; mITT, modified intent-to-treat; NB32, naltrexone 32mg plus bupropion; NB48, naltrexone 48mg plus bupropion
Source: Apovian et al. 2013[17]

COR-BMOD study

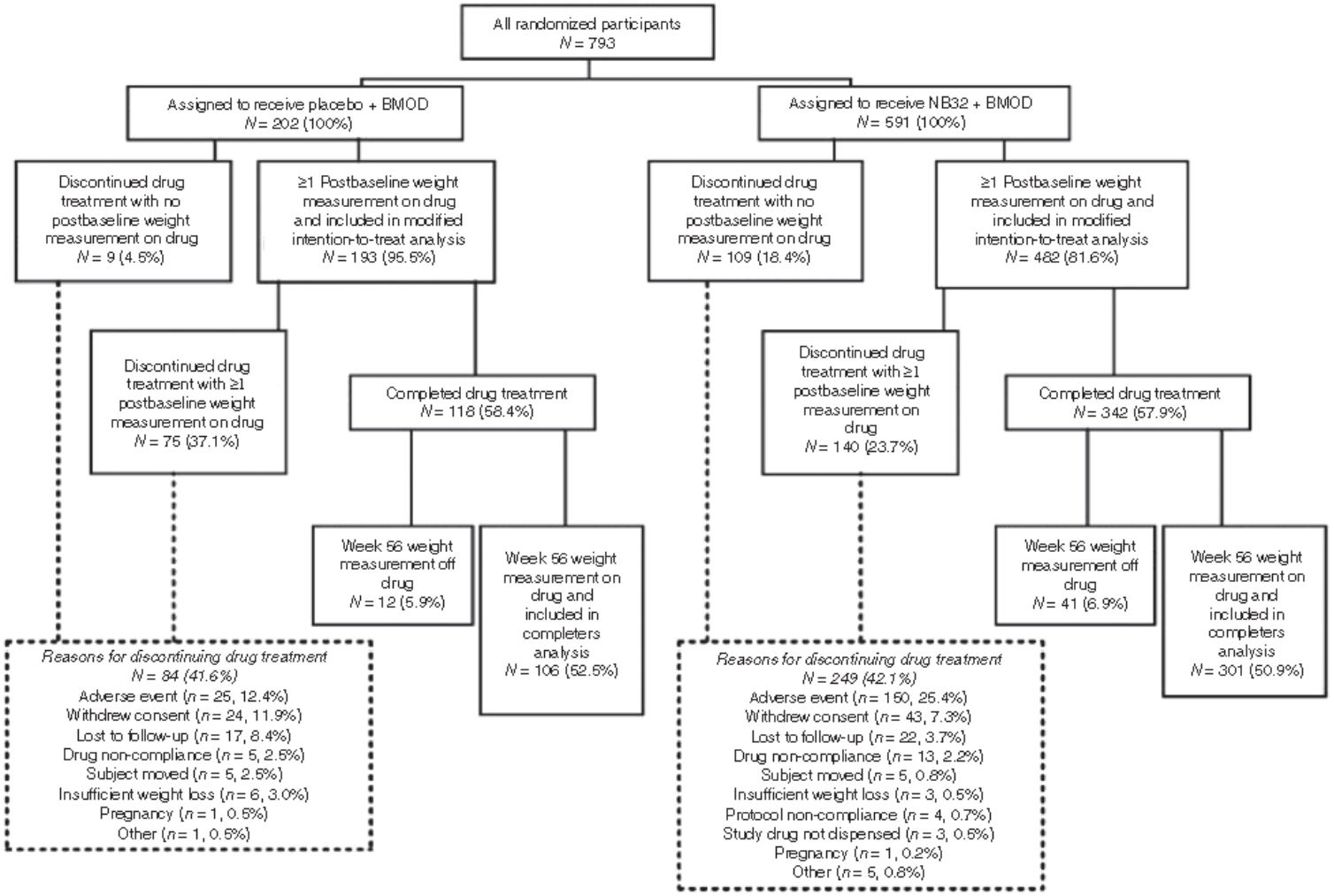
A total of 793 patients were randomised to treatment with NB32 + intensive behaviour modification (BMOD) or placebo + BMOD. Of the 202 patients randomised to placebo + BMOD, 193 (95.5%) qualified for inclusion in the mITT population. A total of 591 patients were randomised to treatment with NB32 + BMOD, of these, 482 (81.6%) qualified for inclusion in the mITT population.[18] During the first 4 weeks of the study, 2.0% of patients in the placebo + BMOD group and 14.0% of patients in the NB32 + BMOD group (p=0.038) did not provide a post baseline measurement of weight on study drug because of study drug discontinuation related to an AE.[18]

Over the 56-week trial, 41.6% of patients in placebo + BMOD discontinued study drug, compared with 42.1% of NB32 + BMOD. A greater percentage of participants who received NB32 + BMOD, compared to patients receiving placebo + BMOD discontinued because of an AE (25.4 vs 12.4%, respectively; p<0.001). By contrast, a greater percentage of patients in the placebo + BMOD group than in NB32 + BMOD discontinued due to withdrawal of consent (11.9 vs 7.3%, respectively; p=0.042), lost to follow-up (8.4 vs 3.7%, respectively; p=0.008), or self-perceived insufficient weight loss (3.0 vs 0.5%, respectively; p=0.004).[18]

A CONSORT flow diagram is presented in Figure 5.

The mean duration of exposure to study drug was 42.62 weeks for the placebo group and 38.63 weeks for the NB32 group. The placebo group represents 161.7 patient years of exposure compared to 427.4 patients-years exposure for the NB32 group.[63] An exposure of ≤4 weeks was observed for 18.2% of patients in the NB32 group compared to 5.0% in the placebo group, consistent with the study drug discontinuation rate observed at Week 4 for the NB32 group over placebo (19.1% vs 5.9%) for the safety analysis set.[63]

Figure 5: CONSORT diagram of participant flow in COR-BMOD



Source: Wadden et al. 2011[18]

COR-DM Study

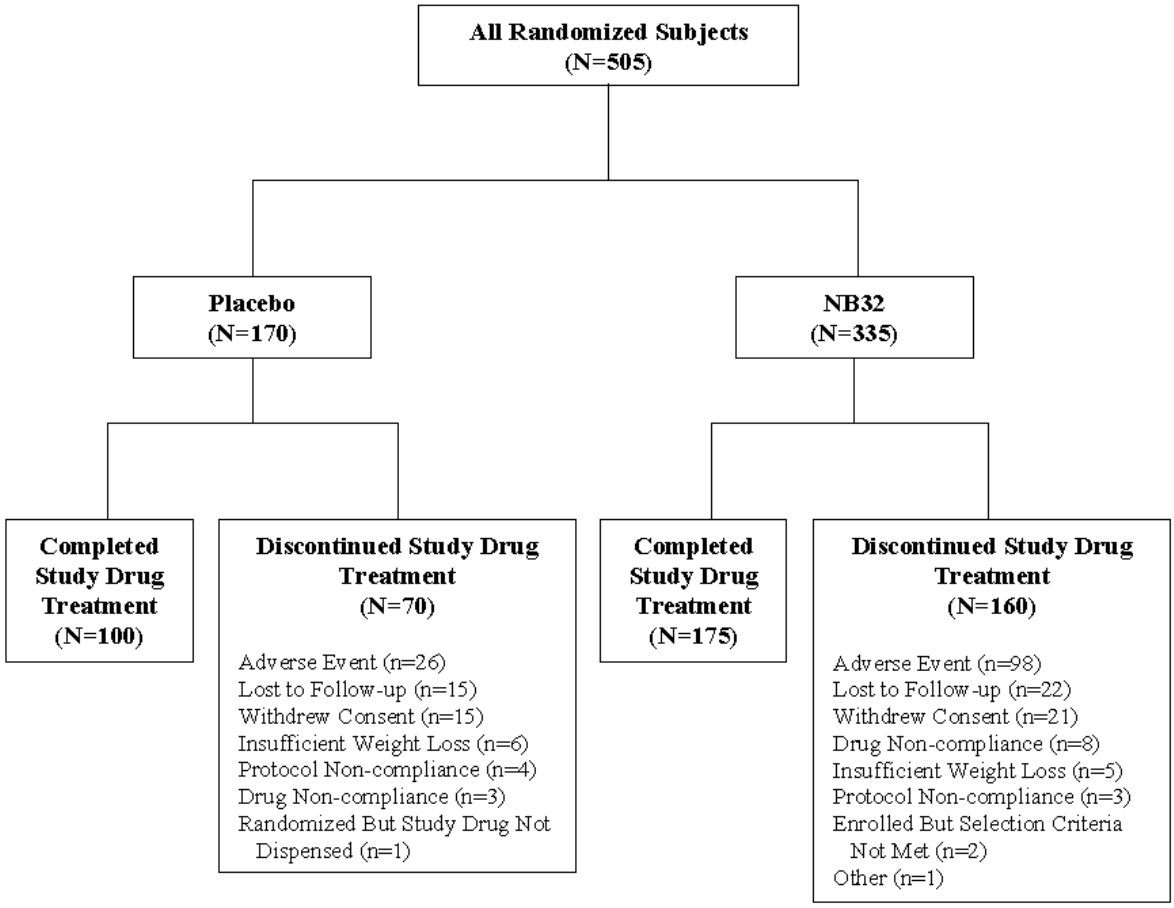
A total of 505 patients were randomised, 335 to the NB32 group and 170 to the placebo group. Of these, 265 (79.1%) and 159 (93.5%) of patients in the NB32 and placebo groups, respectively, were included in the mITT population.[64]

Over the 56-week trial, 47.8% of patients in the NB32 group discontinued the study drug compared with 41.2% in the placebo group. A greater percentage of patients who received NB32 compared with placebo discontinued owing to an AE (29.3 vs 15.3%). Conversely, a greater percentage of patients receiving placebo compared with NB were lost to follow-up (8.8 vs 6.6%), withdrew consent (8.8 vs 6.3%), or withdrew because of self-perceived insufficient weight loss (3.5 vs 1.5%).[19]

A CONSORT flow diagram is presented in Figure 6.

The mean duration of exposure to study drug was 35.07 weeks for the NB32 group and 41.69 weeks for the placebo group. Patient-years of exposure was 22.6 for the NB32 group and 133.7 for the placebo group.[64] An exposure of ≤4 weeks was observed for 22.1% of randomised patients in the NB32 group and 6.5% of patients in the placebo group, consistent with the study drug discontinuation rate observed at Week 4 for the NB32 group (22.1%) and placebo group (7.1%).[64]

Figure 6: CONSORT diagram of participant flow in COR-DM



Source: Orexigen, 2009[64]

Patient characteristics

In all trials, patient demographics and disease characteristics were generally representative of the patient population observed in clinical practice. Patient demographics for all trials are presented in Table 15.

COR-I

Similar to previous clinical trials in obese patients, the patient population in COR-I was predominantly female (85.1%). The mean age of patients was 44.1 years (range: 18 to 66 years), and most patients were white (75.0%). For all randomised patients, mean body weight at baseline was 99.55kg, and mean BMI was 36.17kg/m^2.[16] A substantial proportion of patients had at least one cardiometabolic risk factor at baseline including 49.3% with dyslipidaemia and 20.7% with hypertension[16], and current alcohol use was common (43.2%).[56]

Metabolic syndrome was defined as meeting at least three of the five following criteria at baseline: 1) waist circumference >102cm (men) or 88cm (women); 2) triglycerides ≥150mg/dL; 3) high-density lipoprotein (HDL) cholesterol <40mg/dL (men) or <50mg/dL (women); 4) systolic blood pressure ≥130mmHg and diastolic blood pressure ≥85mmHg; or 5) blood glucose ≥100mg/dL. Impaired fasting glucose was defined as fasting glucose ≥100mg/dL at baseline). A total of 26.4% of subjects had metabolic syndrome, and 25.2% had impaired fasting glucose.[56]

In general, the treatment groups were well-balanced with regard to patient demographics and baseline characteristics, and there were no clinically meaningful differences between the treatment groups for any variable. Demographic characteristics were similar for the mITT population.

COR-II

Again, the study population was predominantly female (84.7%). The mean age of patients was 44.32 years, and most patients were white (83.5%). For all randomised patients, tobacco use and alcohol use was reported by 10.7% and 45.4% respectively.[62]

Mean body weight at baseline was 99.95kg, mean BMI was 36.17kg/m^2, and 58.4% of patients had a BMI ≥35kg/m^2. More than half of all randomised patients had at least one CV risk factor, including 55.0% with dyslipidaemia and 21.3% with hypertension. Additionally, 30.1% of patients had metabolic syndrome, 27.2% had impaired fasting glucose, and the mean (SD) homeostasis model assessment of insulin resistance (HOMA-IR) was 3.53 (3.97).[62]

Similar demographic characteristics were reported for the mITT population. Treatment groups were well balanced with no clinically meaningful difference between the treatment groups for any demographic variable.

COR-BMOD

As before, the study population was predominantly female (89.9%). The mean age of patients was 45.8 years, and most patients were white (69.9%). Current alcohol use was common for all randomised patients (44.3%); however, none of the patients were current tobacco users, consistent with the study entry criterion that prohibited tobacco use for at least 6 months before screening.[18, 63]

Mean body weight at baseline for the study population was 100.60kg, mean BMI was 36.50kg/m^2, and 64.7% of subjects had a BMI ≥35kg/m^2. For all randomised patients, a substantial percentage of subjects had at least one cardio-metabolic risk factor at baseline, including 44.3% with dyslipidaemia and 15.5% with hypertension.[18, 63]

The treatment groups were well-balanced with respect to patient demographics, and there was no clinically meaningful difference between the treatment groups for any demographic variable. Similar demographic characteristics were reported for the mITT population.

COR-DM

The mean age of randomised patients was 53.83 years and 56.4% of patients were female. A total of 89.5% were former or non-smokers. Most patients were white (79.4%). At baseline, mean body weight was 104.51kg, mean waist circumference was 114.11cm, mean BMI was 36.40kg/m^2, and 62.6% had a BMI ≥35kg/m^2. The majority of subjects had at least one cardio-metabolic risk factor at baseline, including dyslipidaemia (84.2%) and hypertension (62.4%).[64]

As this trial included patients with T2DM, the demographic characteristics were different from the previous trials discussed. Namely, patients in this study were of an older patient group, and there was a more even balance between genders. In addition, mean body weight was slightly higher although mean BMI was similar to the other trials. Additionally, a much higher proportion of patients had cardio-metabolic risk factors including hypertension and dyslipidaemia.

The treatment groups were well-balanced with respect to patient demographics, and similar demographic information was reported for the mITT population.

Table 15: Characteristics of participants in the studies across treatment groups (all randomised patients)

COR-I			
	NB32 (n=583)	**NB16 (n=578)**	**Placebo (n=581)**
Age, mean years (SD)	44.4 (11.1)	44.4 (11.3)	43.7 (11.1)
Sex, female, n (%)	496 (85)	490 (85)	496 (85)
Ethnicity, n (%)	White: 440 (75) Black: 106 (18)	White: 427 (74) Black: 122 (21)	White: 440 (76) Black: 110 (19)

	Other: 37 (6)	Other: 29 (5)	Other: 31 (5)
Weight, mean kg (SD)	99.7 (15.9)	99.5 (14.8)	99.5 (14.3)
BMI, mean kg/m^2 (SD)	36.1 (4.4)	36.2 (4.3)	36.2 (4.0)
Smoker, n (%)	65 (11)	56 (10)	65 (11)
Hypertension, n (%)	130 (22)	117 (20)	113 (19)
Dyslipidaemia, n (%)	284 (49)	287 (50)	288 (50)
Alcohol use, n (%)	254 (43.6)	254 (43.9)	244 (42.0)
COR-II			
	NB32 (n=1001)	**Placebo (n=495)**	
Age, mean years (SD)	44.3 (11.2)	44.4 (11.4)	
Sex, female, n (%)	847 (84.6)	420 (84.8)	
Ethnicity, n (%)	White: 835 (83.4) Black: 133 (13.3) Other: 30 (3)	White: 414 (83.6) Black: 72 (!4.5) Other: 20 (2)	
Weight, mean kg (SD)	100.3 (16.6)	99.2 (15.9)	
BMI, mean kg/m^2 (SD)	36.2 (4.5)	36.1 (4.3)	
Smoker, n (%)	108 (10.8)	52 (10.5)	
Hypertension, n (%)	212 (21.2)	106 (21.4)	
Dyslipidaemia, n (%)	560 (55.9)	263 (53.1)	
Alcohol use, n (%)	462 (46.2)	217 (43.8)	
COR-BMOD			
	NB32 (n=591)	**Placebo (n=202)**	
Age, mean years (SD)	45.9 (10.4)	45.6 (11.4)	
Sex, female, n (%)	528 (89.3)	185 (91.6)	
Ethnicity, n (%)	White: 405 (68.5) Black: 145 (24.5) Other: 41 (6.9)	White: 149 (73.8) Black: 44 (21.8) Other: 9 (4.5)	
Weight, mean kg (SD)	100.2 (15.4)	101.9 (15.0)	
BMI, mean kg/m^2 (SD)	36.3 (4.2)	37.0 (4.2)	
Hypertension, n (%)	86 (14.6)	37 (18.3)	
Dyslipidaemia, n (%)	270 (45.7)	81 (40.1)	
Alcohol use, n (%)	251 (42.5)	100 (49.5)	
COR-DM			
	NB32 (n=335)	**Placebo (n=170)**	
Age, mean years (SD)	54.0 (9.1)	53.5 (9.8)	
Sex, female, n (%)	195 (58.2)	90 (52.9)	
Ethnicity, n (%)	White: 261 (77.9)	White: 140 (82.4)	

	Black: 63 (18.8) Other: 11.1 (3.3)	Black: 18 (10.6) Other: 12 (7)
Weight, mean kg (SD)	104.2 (18.9)	105.1 (17.0)
BMI, mean kg/m^2 (SD)	36.4 (4.8)	36.4 (4.5)
Smoker, n (%)	38 (11.3)	15 (8.8)
Hypertension, n (%)	212 (63.3)	103 (60.6)
Dyslipidaemia, n (%)	280 (83.6)	145 (85.3)
Alcohol use, n (%)	96 (28.7)	69 (40.6)
Key: BMI, body mass index; SD, standard deviation. **Source:** Apovian et al. 2013[17]; Greenway et al. 2010[16]; Hollander et al. 2013[19]; Wadden et al. 2011[18]; Orexigen, 2010[62]; Orexigen, 2009[64]; Orexigen, 2010[56]; Orexigen, 2010[63]		

NB-CVOT

Full details of the statistical analysis, participant flow, baseline characteristics and quality assessment are presented in Appendix 3.

4.6 Quality assessment of the relevant randomised controlled trials

All four pivotal RCTs were conducted in line with Good Clinical Practice (GCP) guidelines, with measures taken to reduce the risk of bias.[16-19] All trials are thought to reflect routine clinical practice in England regarding population, comparator choice, treatment administration and outcomes assessed. Outcome assessments were conducted in accordance with trial validated methodology.

A central randomisation system was adopted in all trials. In studies COR-I, COR-II and COR-BMOD, randomisation was stratified by study centre[16-18], while in COR-DM randomisation was stratified by baseline HbA1c levels and sulfonylurea use.[19] All studies were double-blind. Primary efficacy analysis was conducted on the mITT population which included patients with at least one post-baseline assessment while on study medication, and all missing data post baseline were accounted for with the LOCF method. Pre-planned sensitivity analyses including ITT population analysis and additional methods of data imputation substantiated the results of the primary analysis (see Appendix 5).

Quality assessment in accordance with the NICE-recommended checklist for RCT assessment of bias is summarised in Table 16 and presented in full in Appendix 4.

Table 16: Quality assessment results for RCTs

	COR-I	COR-II	COR-BMOD	COR-DM
Was randomisation carried out appropriately?	Yes	Yes	Yes	Yes
Was the concealment of treatment allocation adequate?	Yes	Yes	Yes	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes	Yes	Yes
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Yes	Yes	Yes
Were there any unexpected imbalances in drop-outs between groups?	No	No	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No	No	No
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes, ITT analysis was conducted as part of the sensitivity analyses. The primary analysis set was the mITT population.	Yes, ITT analysis was conducted as part of the sensitivity analyses. The primary analysis set was the mITT population.	Yes, ITT analysis was conducted as part of the sensitivity analyses. The primary analysis set was the mITT population.	Yes, ITT analysis was conducted as part of the sensitivity analyses. The primary analysis set was the mITT population.
How closely do the RCT(s) reflect routine clinical practice	Population, treatment arms, and outcomes all relevant to clinical practice in NHS England.	Population, treatment arms, and outcomes all relevant to clinical practice in NHS England.	Population, treatment arms, and outcomes all relevant to clinical practice in NHS England.	Population, treatment arms, and outcomes all relevant to clinical practice in NHS England.

Key: NHS, National Health Service; RCT, randomised controlled trial.
Source: Apovian et al. 2013[17]; Greenway et al. 2010[16]; Hollander et al. 2013[19]; Wadden et al. 2011[18]

4.7 Clinical effectiveness results of the relevant randomised controlled trials

The data from all four pivotal RCTs demonstrates clear evidence of the clinical benefit of NB32, supporting its use for the management of weight in adult patients who are overweight or obese with one or more weight-related comorbidities.

COR-I study

Co-primary efficacy analysis: mean percent change in body weight

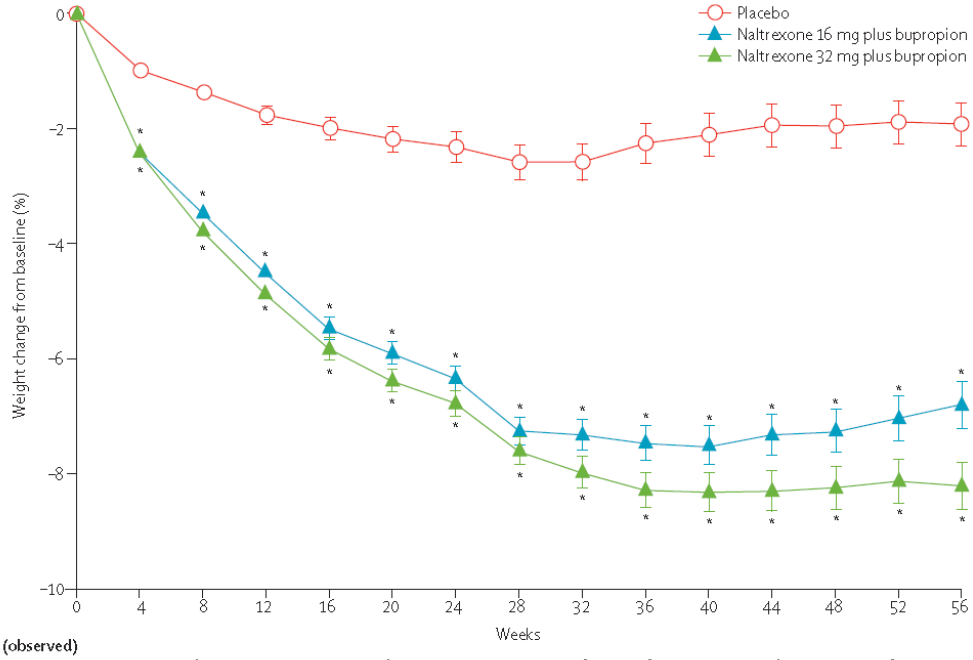
Weight loss in patients assigned to NB began early (Week 4) and was sustained for the duration of the 56-week trial. Maximum weight loss in the combination treatment arms was generally achieved between 28 and 36 weeks. In the primary analysis population (mITT), weight loss was significantly greater in the NB32 (LS mean change in body weight: -6.1%) and NB16 (-5.0%) arms, than in the placebo group (-1.3%) (Table 17). Weight loss in patients who completed 56 weeks of treatment was also greater in the NB32 and NB16 arms (-8.1% and -6.7%, respectively) than in the placebo group (-1.8%), as depicted in Figure 7 (observed data).

Table 17: Percent change in body weight from baseline, COR-I study, mITT population

	NB32 (n=471)	**NB16 (n=471)**	**Placebo (n=511)**
Baseline weight, mean kg (SD)	100.2 (16.3)	100.1 (14.4)	99.3 (14.3)
Week 56 weight, mean kg (SD)	94.2 (17.4)	95.3 (15.8)	98.0 (15.2)
Percent change from baseline at Week 56:			
Mean (SD)	-6.1 (7.1)	-5.0 (6.8)	-1.3 (5.7)
LS Mean (SE)	-6.1 (0.3)	-5.0 (0.3)	-1.3 (0.3)
NB16 or NB32 minus placebo:			
Diff of LS Mean (95% CI)	-4.8 (-5.6, -4.0)	-3.7 (-4.5, -2.8)	
p-value	<0.001	<0.001	
NB32 minus NB16:			
Diff of LS Mean (95% CI)	-1.1 (-2.0, -0.3)		
p-value	0.008		

Key: diff, difference; LS, least squares; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion; SD, standard deviation; SE, standard error.
Source: Greenway et al. 2010[16]; Orexigen, 2010[56]; EMA, 2014[53]

Figure 7: Percentage change from baseline in body weight at each visit during 56 weeks, COR-I study, mITT population (observed data)



Number of participants by visit (observed)														
Placebo	507	463	420	394	365	353	327	318	308	302	296	291	289	277
Naltrexone 16 mg plus bupropion	467	410	373	351	346	341	311	311	302	297	300	284	283	273
Naltrexone 32 mg plus bupropion	467	411	391	372	365	361	343	327	321	316	311	305	298	284

Notes: *, p<0.0001 compared with placebo.
Source: Greenway et al. 2010[16]

Co-primary efficacy analysis: proportion of patients with ≥5% decrease in body weight

A significantly greater proportion of patients in the NB32 and NB16 groups achieved a decrease in bodyweight of ≥5% compared to patients in the placebo group (48% and 39%, respectively, vs 16%; p<0.001 for both).[16] This was seen as early as Week 4.

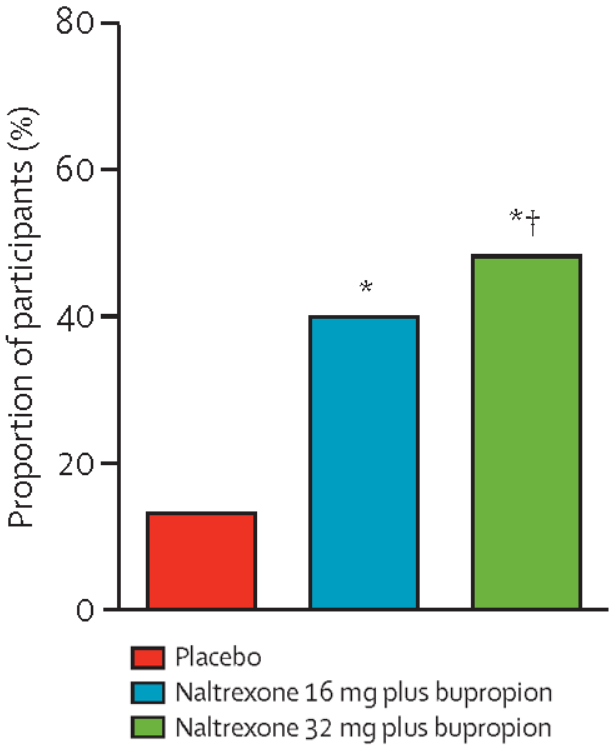
Summary statistics for the proportion of patients with ≥5% decrease in bodyweight from baseline to Week 56 is presented in Table 18 and Figure 8.

Table 18: Patients with ≥5% decrease in bodyweight from baseline to Week 56, COR-I study, mITT population

	NB32 (n=471)	NB16 (n=471)	Placebo (n=511)
Patients with ≥5% decrease in weight, n (%)	226 (48.0)	186 (39.5)	84 (16.4)
95% CI	43.5, 52.5	35.1, 44.0	13.2, 19.7
NB32 or NB16 vs placebo:			
OR (95% CI)	4.9 (3.6, 6.6)	3.4 (2.5, 4.6)	
p-value	<0.001	<0.001	
NB32 vs NB16			
OR (95% CI)	1.4 (1.1, 1.9)		
p-value	<0.01		

Key: CI, confidence interval; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion; OR, odds ratio.
Source: Greenway et al. 2010[16]; Orexigen, 2010[56]; EMA, 2014[53]

Figure 8: Patients with ≥5% decrease in bodyweight from baseline to Week 56, COR-I study, mITT population



Notes: *, p<0.0001 compared with placebo; †, p=0.0099 for naltrexone 32mg plus bupropion compared with naltrexone 16mg plus bupropion.
Source: Greenway et al. 2010[16]

To address potential bias associated with early discontinuation of patients from the study and missing data, sensitivity analyses were performed for the co-primary variables on the ITT, completers and all randomised population. Results of these are presented in Appendix 5. Results of the sensitivity analyses were consistent with results obtained for the primary analyses using the mITT population, indicating significantly greater weight loss, and a greater proportion of patients with ≥5% weight loss, with NB32 compared with placebo, irrespective of the analysis method.

Secondary efficacy analysis

A summary of key secondary endpoints is presented in Table 19.

As seen in the primary efficacy analysis based on ≥5% weight loss, a greater proportion of participants in the NB groups achieved a decrease of ≥10% in bodyweight compared with the placebo group ($p<0.001$ for both observations). In addition, waist circumference was decreased in patients who received NB32 and showed a significant difference compared to placebo ($p<0.001$).

Patients assigned to NB32 showed significant improvements to Week 56 in numerous obesity-associated CV risk factors compared to placebo. LS mean percent change for hs-CRP levels was -29% by Week 56 in patients receiving NB32. In addition, levels of low-density lipoprotein (LDL) cholesterol and triglycerides decreased, while HDL cholesterol levels increased in the NB32 group (vs placebo; $p<0.01$). Furthermore, greater weight loss among patients within the NB treatment groups was associated with larger decreases in systolic and diastolic blood pressure.

In patients with ‘pre-diabetes’, NB32 treatment was also associated with a reduction in diabetic specific risk factors. Fasting insulin levels and insulin resistance were significantly reduced in both NB groups compared to placebo. Similarly, fasting blood glucose was significantly decreased from baseline in the NB32 group compared to placebo ($p=0.01$).

HRQL and PRO measures

A summary of all weight-related quality of life tools used across the four trials, along with example questionnaires, are presented in Appendix 6.

Patients assigned to NB showed greater improvements in the IWQOL-Lite total score than patients assigned to placebo (p<0.001 for both comparisons).[16] These improvements occurred as early as Week 8 and continued throughout the study.[16] In particular, patients receiving NB32 showed greatest improvements in the physical function and self-esteem subscales (mean change of 15.3 and 18.8 respectively; p<0.001 vs placebo for both comparisons).

Greater improvement in eating control for NB32 compared to placebo (p<0.05) was consistently observed in the COE questionnaire at all post-baseline assessments beginning at Week 8 and persisting for the duration of the trial. This indicates that after treatment with NB32, patients show reduced hunger or desire for sweet, non-sweet, or starchy foods; increased feeling of fullness; reduced incidence and strength of food cravings; reduced eating in response to food cravings; and increased ability to resist food cravings and control eating.[16] In the sweets and carbohydrate subscales of the FCI minimal differences were seen between the NB32 and placebo groups.

Furthermore, treatment with NB32 resulted in a mean change of -0.4 in the IDS-SR score compared to -0.6 in placebo-treated patients, demonstrating that NB32 does not increase risk of suicide and depressive behaviours.

The significant treatment effects on obesity-related metabolic and quality of life parameters observed in this study support the clinical relevance of the treatment effects observed on total body weight.

Table 19: Summary of key secondary endpoints, COR-I study, mITT population

	NB32 (n=471)	**NB16 (n=471)**	**Placebo (n=511)**
Proportion of patients with ≥10% decrease in body weight			
n (%)	116 (24.6)	95 (20.2)	38 (7.4)
OR (95% CI) [p-value]	4.2 (2.8, 6.2) [<0.001]	3.2 (2.1, 4.8) [<0.001]	
Change in waist circumference (cm)			
n	356	342	348
Baseline, mean (SD)	108.8 (11.3)	109.9 (11.2)	110.0 (12.2)
Week 56, mean (SD)	102.6 (12.4)	104.7 (12.7)	107.4 (12.9)
Mean change (SD)	-6.3 (8.2)	-5.2 (8.5)	-2.6 (7.1)
LS mean (SE)	-6.2 (0.4)	-5.0 (0.4)	-2.5 (0.4)

	NB32 (n=471)	NB16 (n=471)	Placebo (n=511)
Diff of LS mean (95% CI)	-3.8 (-4.9, -2.6)	-2.6 (-3.7, -1.4)	
p-value	<0.001	<0.001	
Change in fasting insulin levels (µIU/mL)			
n	344	309	326
Baseline, geometric mean	11.1	11.4	11.3
Week 56, geometric mean	9.5	10.2	11.0
Percent change from baseline:			
LS percent change	-17.1	-11.9	-4.6
p-value	0.001	0.063	
Change in fasting HDL cholesterol levels (mg/dL)			
n	359	333	345
Baseline, mean (SD)	51.9 (13.6)	52.3 (13.4)	52.0 (13.6)
Week 56, mean (SD)	55.3 (14.2)	55.6 (14.5)	51.9 (13.8)
Mean change (SD)	3.4 (8.8)	3.3 (8.7)	-0.2 (7.8)
LS mean (SE)	3.4 (0.5)	3.4 (0.5)	-0.1 (0.5)
Diff of LS mean (95% CI)	3.5 (2.3, 4.7)	3.4 (2.2, 4.7)	
p-value	<0.001	<0.001	
Change in fasting LDL cholesterol levels (mg/dL)			
n	358	332	345
Baseline, mean (SD)	118.8 (32.6)	124.7 (32.5)	119.7 (34.8)
Week 56, mean (SD)	115.5 (32.4)	120.5 (31.5)	117.2 (34.0)
Mean change (SD)	-3.3 (22.3)	-4.2 (21.8)	-2.5 (24.1)
LS mean (SE)	-4.4 (1.2)	-3.7 (1.2)	-3.3 (1.2)
Diff of LS mean (95% CI)	-1.1 (-4.3, 2.0)	0.4 (-3.6, 2.8)	
p-value	0.484	0.811	
Change in fasting triglycerides (mg/dL)			
n	359	333	345
Baseline, geometric mean	116.0	118	113.2
Week 56, geometric mean	102.6	109.3	111.8
Percent change from baseline:			
LS percent change	-12.7	-8.0	-3.1
p-value	<0.001	0.046	
Change in hs-CRP levels (mg/L)			
n	353	331	340

	NB32 (n=471)	NB16 (n=471)	Placebo (n=511)
Baseline, geometric mean (SD)	3.8 (2.8)	3.9 (2.6)	3.6 (2.8)
Week 56, geometric mean	2.8	2.8	3.1
Percent change from baseline:			
LS percent change (95% CI)	-29.0 (-34.8, -22.7)	-28.0 (-34.1, -21.4)	-16.7 (-23.7, -9.0)
p-value	0.008	0.016	
Change in fasting blood glucose levels (mg/dL)			
n	361	336	348
Baseline, mean (SD)	94.2 (12.1)	95.2 (11.5)	93.9 (11.2)
Week 56, mean (SD)	91.4 (11.4)	92.8 (12.3)	93.2 (11.2)
Mean change (SD)	-2.8 (12.2)	-2.4 (11.6)	-0.7 (10.5)
LS mean (SE)	-3.2 (0.6)	-2.4 (0.6)	-1.3 (0.6)
Diff of LS mean (95% CI)	-1.9 (-3.4, -0.5)	-1.1 (-2.6, 0.4)	
p-value	0.01	N/A	
Change in HOMA-IR levels			
n	341	305	325
Baseline, geometric mean (SD)	2.6 (2.0)	2.6 (2.0)	2.6 (2.0)
Week 56, geometric mean	2.1	2.3	2.5
Percent change from baseline:			
LS percent change (95% CI)	-20.2 (-25.3, -14.8)	-14.3 (-20.1, -8.1)	-5.9 (-12.1, 0.7)
p-value	0.0003	0.044	
Change in systolic blood pressure (mmHg)			
Baseline, mean (SD)	118.9 (9.8)	119.5 (9.9)	119.0 (9.8)
Week 56, mean (SD)	118.7 (11.1)	119.4 (11.2)	116.9 (10.2)
Mean change (SD)	-0.2 (9.8)	-0.0 (9.4)	-2.1 (9.6)
LS mean (SE)	-0.1 (0.4)	0.3 (0.4)	-1.9 (0.4)
Diff of LS mean (95% CI)	1.8 (0.8, 2.9)	2.2 (1.2, 3.3)	
p-value	<0.001	<0.001	
Change in diastolic blood pressure (mmHg)			
Baseline, mean (SD)	77.1 (7.2)	76.6 (7.2)	77.3 (6.6)
Week 56, mean (SD)	76.7 (7.5)	76.5 (7.8)	75.9 (7.3)
Mean change (SD)	-0.4 (7.2)	-0.1 (6.9)	-1.4 (6.7)
LS mean (SE)	0.0 (0.3)	0.1 (0.3)	-0.9 (0.3)

	NB32 (n=471)	NB16 (n=471)	Placebo (n=511)
Diff of LS mean (95% CI)	0.9 (0.1, 1.7)	1.0 (0.2, 1.7)	
p-value	0.022	0.015	
Change in IWQOL-Lite total scores			
n	417	422	468
Baseline, mean (SD)	70.3 (16.5)	70.7 (17.0)	71.8 (17.2)
Week 56, mean (SD)	83.3 (14.7)	82.5 (14.7)	80.1 (15.5)
Mean change (SD)	13.0 (12.8)	11.8 (12.4)	8.3 (12.1)
LS mean (SE)	12.7 (0.5)	11.7 (0.5)	8.6 (0.5)
Diff of LS mean change (95% CI)	4.1 (2.7, 5.6)	3.1 (1.7, 4.5)	
p-value	<0.001	<0.001	
Change in 21-item COE scores, item #19			
n	409	410	453
Baseline, mean (SD)	58.4 (25.3)	61.3 (22.7)	57.6 (25.5)
Week 56, mean (SD)	44.4 (23.6)	47.5 (22.7)	50.1 (23.0)
Mean change (SD)	-14.0 (27.9)	-13.7 (26.8)	-7.4 (26.0)
LS mean (SE)	-14.5 (1.1)	-12.5 (1.1)	-8.7 (1.0)
Diff of LS mean (95% CI)	-5.8 (-8.7, -3.0)	-3.8 (-6.7, -0.9)	
p-value	<0.001	N/A	
Change in IDS-SR Total score			
N	470	471	511
Baseline, mean (SD)	6.7 (5.5)	6.5 (5.5)	6.2 (5.0)
Week 56, mean (SD)	6.4 (5.1)	6.5 (5.3)	5.6 (4.9)
Mean change (SD)	-0.4 (5.0)	0.0 (5.4)	-0.6 (4.8)
LS mean (SE)	-0.3 (0.2)	0.0 (0.2)	-0.7 (0.2)
Diff of LS mean (95% CI)	0.5 (-0.1, 1.0)	0.7 (0.2, 1.3)	
p-value	0.102	0.008	

Key: CI, confidence interval; COE, control of eating; HDL, high density lipoprotein; HOMA-IR, homoeostasis model-insulin resistance; hs-CRP, high-sensitivity C-reactive protein; IWQOL-Lite, impact of weight on quality of life-lite; LDL, low density lipoprotein; LS, least squares; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion; OR, odds ratio; SD, standard deviation; SE, standard error.
Source: Greenway et al. 2010[16]; Orexigen, 2010[56]; EMA, 2014[53]

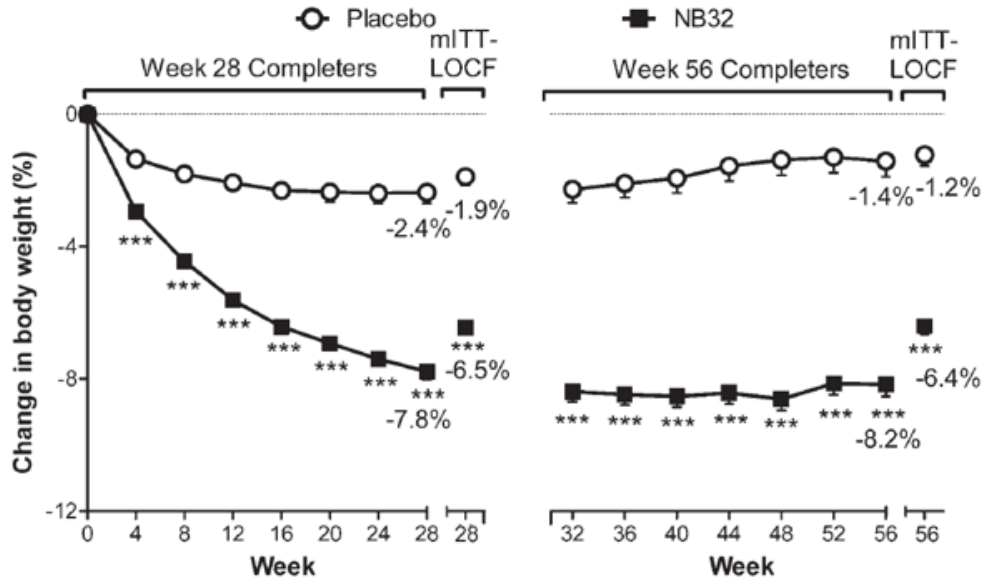
A body composition sub-study was also conducted for this trial, full details of which are provided in Appendix 5.

COR-II study

Co-primary efficacy analysis: mean percent change in body weight

At Week 28, LS mean weight loss was significantly greater for patients treated with NB32 compared to placebo (-6.5% vs -1.9%, respectively; p<0.001 [Figure 9; note that 56 week results are a secondary outcome; Table 20]) in the primary population (mITT).[17] This was continued throughout double-blind treatment to Week 56 in the NB32 groups (Table 21).[17] In the completers population, LS mean percent change at Week 28 was -7.8% in NB32-treated patients compared to -2.4% in placebo-treated patients (p<0.001).

Figure 9: Percent change in bodyweight from baseline, COR-II study, Week 28 and 56 completers and mITT population



Key: LOCF, last observation carried forward; mITT, modified intent-to-treat.
Notes: ***, p<0.001 for NB32 vs placebo; values shown are LS mean ± SE.
Source: Apovian et al. 2013[17]

Co-primary efficacy analysis: proportion of patients with ≥5% decrease in body weight

At Week 28, the proportion of patients with ≥5% decrease in body weight from baseline in the NB32 group was statistically greater compared with placebo-treated patients (55.6% vs 17.5%; p<0.001 [Table 20]) (mITT population). This was observed as early as Week 4.[17]

Table 20: Co-primary efficacy analysis, COR-II study, mITT population

	NB32 (n=825)	Placebo (n=456)
Mean change in body weight		
Baseline weight, mean kg (SD)	100.7 (16.7)	99.3 (16.0)
Week 28 weight, mean kg (SD)	94.2 (17.6)	97.2 (16.2)
Percent change from baseline:		
Mean (SD)	-6.6 (6.1)	-2.1 (4.7)
LS Mean (SE)	-6.5 (0.2)	-1.9 (0.3)
Diff of LS Mean (95% CI)	-4.6 (-5.2, -3.9)	
p-value	<0.001	
Proportion of patients with ≥5% decrease in body weight		
n (%)	459 (55.6)	80 (17.5)
95% CI	52.3, 59.0	14.1, 21.0
OR (95% CI)	6.6 (5.0, 8.8)	
p-value	<0.001	

Key: CI, confidence interval; diff, difference; LS, least squares; NB32, naltrexone 32mg plus bupropion; SD, standard deviation; SE, standard error.
Source: Apovian et al. 2013[17]; Orexigen, 2010[62]; EMA, 2014[53]

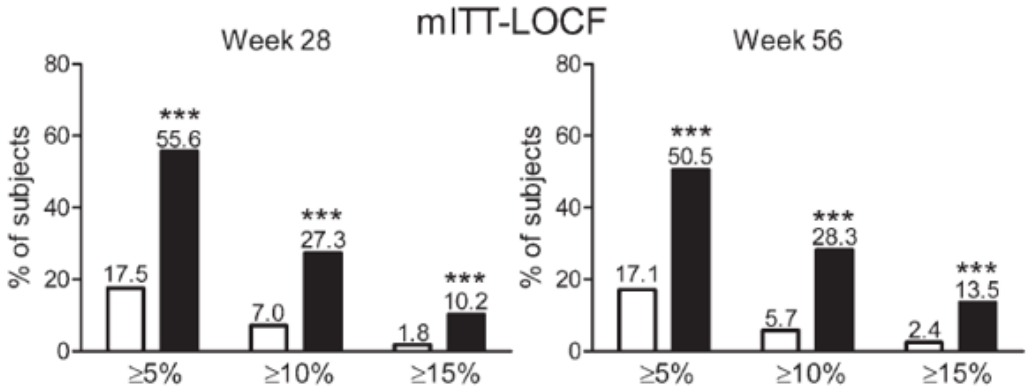
As in the COR-I study, sensitivity analyses were performed to address the potential bias associated with early discontinuations of patients and missing data. Results of these analyses are presented in Appendix 5. Results were consistent with those obtained for the primary analysis using the mITT population, indicating that NB32-treated patients had significantly greater weight loss and a greater proportion of patients with ≥5% weight loss irrespective of the analysis method.

Secondary efficacy analysis

A summary of results for key secondary efficacy analyses are presented in Table 21.

Percentage change in body weight continued to decrease to Week 56, where a statistically significantly greater decrease was observed with NB32 compared to placebo (p<0.001 [Figure 9]) (mITT).[17] In addition, NB32 was associated with a significantly larger proportion of patients achieving ≥10% and ≥15% weight loss compared to placebo at Weeks 28 and 56 (Figure 10).[17]

Figure 10: Proportion of patients with ≥10% and ≥15% weight loss, COR-II study, mITT population



Key: LOCF, last observation carried forward; mITT, modified intent-to-treat.
Notes: ***, p<0.001 for NB32 vs placebo.
Source: Apovian et al. 2013[17]

NB32 also resulted in improvements in various cardiometabolic parameters, including reductions in waist circumference, triglycerides and LDL cholesterol, and increased HDL cholesterol compared to placebo (Table 21). NB32 was also associated with reduced fasting insulin and HOMA-IR, representative of a reduced risk of diabetes development. In most cases, improvements in secondary endpoints were maintained at Week 56.[17]

HRQL and PRO measures

At Week 28, NB32 was associated with improvement in total IWQOL-Lite score (as described previously) compared to placebo (p<0.001). In particular, greater improvements for NB32 compared to placebo were observed in the physical function, self-esteem, and sexual life subscales (p<0.01). These improvements were maintained through Week 56.[17]

Item #19 of the COE questionnaire showed a greater decrease in patients assigned to NB32 (-18.32) compared to placebo-treated patients (-11.09) demonstrating an association between NB32 and improved control of eating and food craving. For both the sweets and carbohydrates subscale scores for the FCI, decreases were observed in both the NB32 and placebo groups, further demonstrating a reduced food craving, as expected based on the innovative mechanism of action of NB32. At

Week 28, a small decrease from baseline in the IDS-SR total score was also observed for both treatment groups, supporting the view that NB32 has no effect on suicide risk of depressive episodes.

Table 21: Summary of key secondary endpoints, COR-II study, mITT population

	NB32 (n=702)	**Placebo (n=456)**
Percent change in body weight to Week 56		
Baseline, mean (SD)	100.2 (16.4)	99.3 (16.0)
Week 56, mean (SD)	93.0 (17.8)	97.9 (16.4)
Mean change (SD)	-7.4 (7.4)	-1.4 (4.9)
LS Mean (SE)	-6.4 (0.3)	-1.2 (0.3)
Diff of LS Mean (95% CI)	-5.2 (-6.0, -4.4)	
p-value	<0.001	
Proportion of patients with ≥5% decrease in body weight to Week 56		
n (%)	355 (50.5)	78 (17.1)
OR (95% CI) [p-value]	5.5 (4.1, 7.5) [p<0.001]	
Proportion of patients with ≥10% decrease in body weight to Week 28		
N	825	456
n (%)	225 (27.3)	32 (7.0)
OR (95% CI) [p-value]	5.4 (3.6, 8.0) [p<0.001]	
Change in waist circumference (cm)		
N	622	315
Baseline, mean (SD)	109.3 (11.9)	108.9 (11.7)
Week 28, mean (SD)	103.0 (12.9)	106.0 (12.1)
Mean change (SD)	-6.3 (7.1)	-2.9 (5.7)
LS Mean (SE)	-6.2 (0.3)	-2.7 (0.4)
Diff of LS Mean (95% CI)	-3.4 (-4.3, -2.5)	
p-value	<0.001	
Change in fasting insulin levels (µIU/mL)		
N	589	286
Baseline, geometric mean (SD)	11.4 (1.9)	10.7 (1.9)
Week 28, geometric mean	9.7	10.8
LS mean percent change (95% CI)	-14.1 (-17.9, -10.2)	-0.5 (-6.5, 5.9)
p-value	<0.001	

	NB32 (n=702)	Placebo (n=456)
Change in fasting HDL cholesterol levels (mg/dL)		
N	625	308
Baseline, mean (SD)	51.4 (13.3)	51.4 (13.1)
Week 28, mean (SD)	52.6 (13.2)	50.2 (12.6)
Mean change (SD)	-1.2 (7.8)	-1.3 (7.5)
LS Mean (SE)	1.2 (0.3)	-1.4 (0.4)
Diff of LS Mean (95% CI)	2.6 (1.6, 3.6)	
p-value	<0.001	
Change in fasting LDL cholesterol levels (mg/dL)		
N	620	308
Baseline, mean (SD)	119.8 (30.2)	117.1 (32.6)
Week 28, mean (SD)	115.4 (31.0)	117.9 (32.4)
LS mean change (SE)	-4.4 (0.9)	0.0 (1.3)
p-value	0.004	
Change in fasting triglycerides (mg/dL)		
N	625	308
Baseline, geometric mean	119.0	113.4
Week 28, geometric mean	110.5	113.3
Percent change from baseline:		
LS Percent change (95% CI)	-7.3 (-9.8, -4.8)	-1.4 (-5.0, 2.4)
p-value	0.007	
Change in hs-CRP levels (mg/L)		
N	607	304
Baseline, geometric mean (SD)	3.9 (2.8)	3.7 (2.7)
Week 28, geometric mean	3.5	3.6
LS mean percent change (95% CI)	-9.4 (-14.8, -3.6)	-1.1 (-9.1, 7.5)
p-value	0.091	
Change in fasting blood glucose levels (mg/dL)		
N	628	310
Baseline, mean (SD)	94.8 (11.2)	94.2 (10.4)
Week 28, mean (SD)	92.6 (10.4)	92.6 (11.0)
LS mean change (SE)	-2.1 (0.4)	-1.7 (0.5)
p-value	0.544	
Change in HOMA-IR levels		
N	580	278

	NB32 (n=702)	Placebo (n=456)
Baseline, geometric mean (SD)	2.7 (2.0)	2.5 (2.0)
Week 28, geometric mean	2.2	2.4
LS mean percent change (95% CI)	-16.4 (-20.4, -12.3)	-4.2 (-10.4, 2.6)
p-value	<0.001	
Change in systolic blood pressure (mmHg)		
N	824	456
Baseline, mean (SD)	118.1 (10.0)	118.2 (10.5)
Week 28, mean (SD)	117.2 (11.5)	116.8 (11.3)
LS mean change (SE)	-0.9 (0.3)	-1.2 (0.4)
p-value	0.556	
Change in diastolic blood pressure (mmHg)		
N	824	456
Baseline, mean (SD)	76.8 (7.0)	76.8 (7.0)
Week 28, mean (SD)	76.9 (7.7)	76.0 (7.6)
LS mean change (SE)	0.2 (0.2)	-0.7 (0.3)
p-value	0.017	
Change in IWQOL-Lite total scores		
N	628	317
Baseline, mean (SD)	72.0 (17.4)	72.9 (15.7)
Week 28, LS mean (SE)	9.9 (0.4)	6.2 (0.6)
Diff of LS mean (95% CI)	3.8 (2.5, 5.1)	
p-value	<0.001	
Change in 21-item COE scores, item #19		
Baseline, mean (SD)	61.9 (24.1)	62.0 (23.5)
Week 28, mean (SD)	48.1 (19.4)	52.2 (19.7)
LS mean change (SE)	-18.3 (0.9)	-11.1 (1.1)
p-value	<0.001	
Change in IDS-SR total scores		
Baseline, mean (SD)	7.2 (6.0)	6.9 95.3)
Week 28, mean (SD)	6.9 (5.3)	6.8 (6.4)
LS mean change (SE)	-0.2 (0.2)	-0.3 (0.2)
p-value	0.844	

	NB32 (n=702)	Placebo (n=456)
Key: CI, confidence interval; COE, control of eating; HDL, high density lipoprotein; HOMA-IR, homoeostasis model-insulin resistance; hs-CRP, high-sensitivity C-reactive protein; IDS-SR, Inventory of Depressive Symptoms - Subject Rated; IWQOL-Lite, impact of weight on quality of life-lite; LDL, low density lipoprotein; LS, least squares; mITT, modified intent-to-treat; NB32, naltrexone 32mg plus bupropion; OR, odds ratio; SD, standard deviation; SE, standard error. **Source:** Apovian et al. 2013[17]; Orexigen, 2010[62]; EMA, 2014[53]		

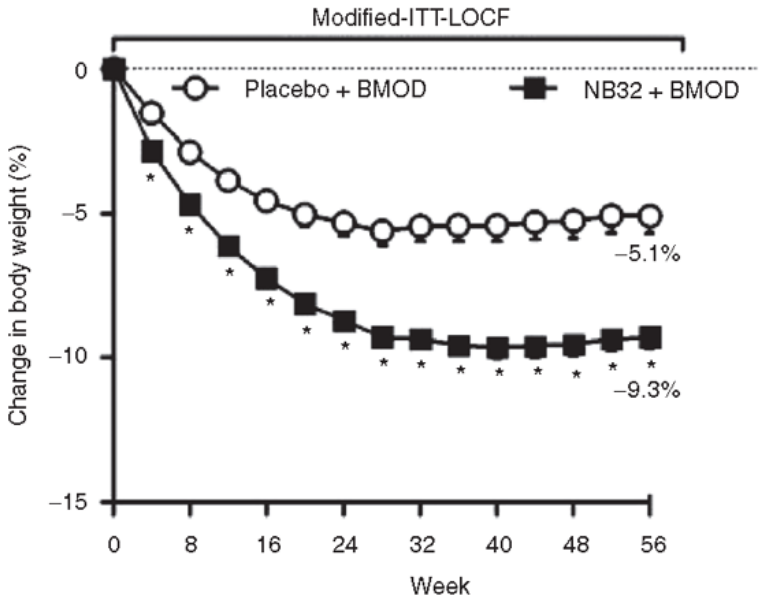
A blood pressure sub-study was also conducted in this trial. Full details of this are provided in Appendix 5.

COR-BMOD study

Co-primary efficacy analysis: mean percent change in body weight

At Week 56, patients in the primary population (mITT) who received NB32 + BMOD achieved a LS mean weight loss of 9.3% compared to only 5.1% for patients treated with placebo + BMOD (p<0.001).[18] This was observed as early as Week 4 (Figure 11).

Figure 11: Percent change in body weight from baseline, COR-BMOD study, mITT population

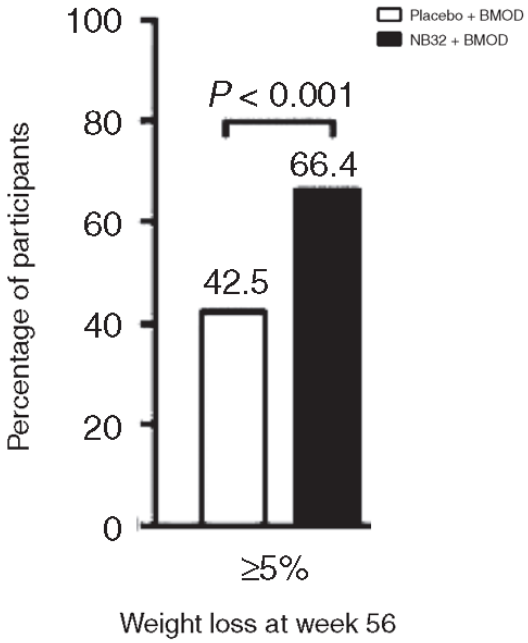


Key: BMOD, behaviour modification; ITT, intent-to-treat; LOCF, last observation carried forward; mITT, modified intent-to-treat; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion.
Notes: *, p<0.001
Source: Wadden et al. 2010[18]

Co-primary efficacy analysis: proportion of patients with ≥5% decrease in body weight

The proportion of patients who achieved ≥5% reduction in baseline weight were greater with NB32 + BMOD than with placebo + BMOD (p<0.001) (Table 22) (mITT population). More than 1.5 times as many patients who received NB32 + BMOD achieved this endpoint compared with placebo + BMOD patients (Figure 12).[18]

Figure 12: Percentage of patients losing ≥5% body weight, COR-BMOD study, mITT population



Key: BMOD, behaviour modification; mITT, modified intent-to-treat
Source: Wadden et al. 2010[18]

Table 22: Co-primary efficacy analysis, COR-BMOD study, mITT population

	NB32 + BMOD (n=482)	Placebo + BMOD (n=193)
Mean change in body weight		
Baseline weight, mean kg (SD)	100.7 (15.4)	101.9 (15.0)
Week 56 weight, mean kg (SD)	91.0 (17.1)	96.4 (17.1)
Percent change from baseline:		
Mean (SD)	-9.7 (8.5)	-5.5 (7.9)

	NB32 + BMOD (n=482)	Placebo + BMOD (n=193)
LS Mean (SE)	-9.3 (0.4)	-5.1 (0.6)
Diff of LS Mean (95% CI)	-4.2 (-5.6, -2.9)	
p-value	<0.001	
Proportion of patients with ≥5% decrease in body weight		
N (%)	320 (66.4)	82 (42.5)
95% CI	62.2, 70.6	35.5, 49.5
OR (95% CI)	2.9 (2.0, 4.1)	
p-value	<0.001	

Key: BMOD, behaviour modification; CI, confidence interval; diff, difference; LS, least squares; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion; SD, standard deviation; SE, standard error.
Source: Wadden et al. 2010[18]; Orexigen, 2010[63]; EMA, 2014[53]

As in the COR-I and COR-II studies, sensitivity analyses were conducted and the results of these were generally consistent with those obtained for the primary analysis. However, it should be noted that in the completers analysis set, patients who received NB32 showed a mean reduction of 11.5%, compared to the reduction of 9.7% seen for the mITT population.[63] In addition, post-hoc sensitivity analyses were conducted using the ITT analysis set and again, results were consistent with those obtained for the primary analysis.

Secondary efficacy analysis

A summary of key secondary endpoints is presented in Table 23.

The proportion of patients with ≥10% decrease in body weight was significantly higher in the NB32 group compared to placebo (41.5% vs 20.2%, respectively; p<0.001).[53]

Waist circumference declined significantly (p<0.001) more at Week 56 with NB32 + BMOD than with placebo + BMOD, as did plasma triglycerides (p=0.004), insulin (p=0.003), and HOMA-IR (p=0.002). HDL cholesterol increased significantly (p<0.001) more with NB32 + BMOD than with placebo + BMOD. Mean LDL cholesterol levels increased slightly in both treatment groups, with a difference of -2.70mg/dL between the NB32 and placebo groups.[63] In addition, the LS percent change in fasting insulin was statistically significantly superior for NB32 compared to

placebo (-28.0% vs -15.5% [p=0.003]).[63] There were no statistically significant differences between groups in changes in hs-CRP.[18]

HRQL and PRO measures

Overall weight-related quality of life, as measured by the IWQOL-Lite total score improved significantly more at all assessment visits with NB32 + BMOD than with placebo + BMOD ($p<0.05$ for all comparisons). In exploratory analyses, patients in with NB32 + BMOD group also reported greater improvements on the physical function and self-esteem subscales than did placebo + BMOD treated patients.[18]

Unlike the COR-I and COR-II studies, item #19 of the COE questionnaire was not specifically defined as a secondary outcome; instead, all questions of the COE questionnaire were pre-defined secondary outcomes in this study. Scores for items measuring hunger, food craving strength, and eating control decrease for both treatment groups to Week 56, with numerically greater decreases for NB32 compared to placebo, suggestive of treatment effects.[63] This included questions such as ‘how hungry have you felt?’, ‘how strong was your desire to eat tasty foods that are not sweet?’, ‘how strong have your food cravings been?’ and ‘how often have you eaten in response to food cravings?’ in addition to item #19 of the questionnaire ‘generally, how difficult has it been to control your eating?’. A similar trend was observed for items measuring mood and alertness.[63]

The LS mean change in IDS-SR score from baseline was similar between the two treatment groups, with a magnitude of treatment difference of 0.09 demonstrating that NB32 + BMOD treatment does not result in an increase in depressive symptoms. Furthermore, mean change in the FCI sweets and carbohydrates subscale score was -2.6 and -2.2 respectively showing an improvement in food cravings, associated with the mechanism of action of NB32.

Table 23: Summary of key secondary endpoints, COR-BMOD study, mITT population

	NB32 + BMOD (n=482)	Placebo + BMOD (n=193)
Proportion of patients with ≥10% decrease in body weight		
n (%)	200 (41.5)	39 (20.2)
OR (95% CI) [p-value]	2.9 (2.0, 4.4) [<0.001]	
Change in waist circumference (cm)		
Baseline, mean (SD)	109.3 (11.4)	109.0 (11.8)
Week 56, mean (SD)	99.1 (12.8)	102.0 (13.1)
LS mean change (95% CI)	-10.0 (-10.9, -9.0)	-6.8 (-8.3, -5.3)
p-value	<0.001	
LS mean percent change (95% CI)	-9.1 (-9.9, -8.2)	-6.1 (-7.5, -4.7)
p-value	<0.001	
Change in fasting insulin levels (µIU/mL)		
Baseline, geometric mean	11.3 (1.8)	11.0 (1.7)
Week 56, geometric mean	7.8 (2.1)	8.8 (1.8)
LS percent change (95% CI)	-28.0 (-32.4, -23.3)	-15.5 (-23.3, -6.8)
p-value	0.003	
Change in fasting HDL cholesterol levels (mg/dL)		
Baseline, mean (SD)	53.6 (13.5)	55.3 (12.9)
Week 56, mean (SD)	58.5 (14.1)	56.9 (13.4)
LS mean change (95% CI)	4.1 (3.1, 5.1)	0.9 (-0.7, 2.4)
p-value	<0.001	
LS percent change (95% CI)	9.4 (7.4, 11.4)	2.8 (-0.3, 6.0)
p-value	<0.001	
Change in fasting LDL cholesterol levels (mg/dL)		
Baseline, mean (SD)	109.5 (27.5)	109.2 (27.3)
Week 56, mean (SD)	115.0 (30.9)	117.3 (33.2)
LS mean change (95% CI)	5.4 (2.8, 8.1)	8.1 (4.0, 12.3)
p-value	0.245	
LS percent change (95% CI)	7.1 (4.3, 9.8)	10.0 (5.7, 14.3)
p-value	0.219	
Change in fasting triglycerides (mg/dL)		
Baseline, geometric mean	111.6 (1.6)	104.6 (1.6)
Week 56, geometric mean	91.4 (1.6)	95.6 (1.6)

	NB32 + BMOD (n=482)	Placebo + BMOD (n=193)
LS percent change (95% CI)	-16.6 (-19.7, -13.5)	-8.5 (-13.7, -3.0)
p-value	0.004	
Change in hs-CRP levels (mg/L)		
Baseline, geometric mean	3.9 (2.7)	4.2 (2.6)
Week 56, geometric mean	2.7 (3.1)	3.1 (3.4)
LS percent change (95% CI)	-25.9 (-32.6, -18.5)	-16.9 (-28.3, -3.7)
p-value	0.165	
Change in fasting blood glucose levels (mg/dL)		
Baseline, mean (SD)	92.4 (10.7)	94.1 (20.1)
Week 56, mean (SD)	90.0 (11.2)	91.6 (14.0)
LS mean change (95% CI)	-2.4 (-3.6, -1.2)	-1.1 (-3.0, 0.8)
p-value	0.225	
LS percent change (95% CI)	-1.5 (-2.9, -0.2)	0.0 (-2.1, 2.1)
p-value	0.185	
Change in HOMA-IR levels		
Baseline, geometric mean	2.6 (1.9)	2.5 (1.8)
Week 56, geometric mean	1.7 (2.2)	2.0 (1.9)
LS percent change (95% CI)	-29.9 (-34.6, -24.9)	-16.6 (-25.0, -7.1)
p-value	0.003	
Change in systolic blood pressure (mm Hg)		
Baseline, mean (SD)	116.9 (9.9)	116.7 (10.9)
Week 56, mean (SD)	115.8 (11.9)	113.0 (11.8)
Mean change (SD)	-1.2 (10.5)	-3.7 (9.1)
LS mean (SE)	-1.3 (0.5)	-3.9 (0.7)
Diff of LS mean (95% CI)	2.6 (1.0, 4.1)	
p-value	0.002	
Change in diastolic blood pressure (mm Hg)		
Baseline, mean (SD)	78.2 (7.2)	77.2 (7.4)
Week 56, mean (SD)	84.3 (7.2)	81.7 (7.0)
Mean change (SD)	6.2 (6.6)	4.5 (6.0)
LS mean (SE)	6.2 (0.3)	4.2 (0.4)
Diff of LS Mean (95% CI)	1.99 (1.1, 2.9)	
p-value	<0.001	
Change in IWQOL-Lite total scores		
Baseline, mean (SD)	71.9 (15.4)	73.8 (15.6)

	NB32 + BMOD (n=482)	Placebo + BMOD (n=193)
Week 56, mean (SD)	85.6 (14.0)	83.7 (14.8)
LS mean change (95% CI)	13.4 (12.3, 14.5)	10.3 (8.6, 12.0)
p-value	<0.001	
LS percent change (95% CI)	23.9 (22.0, 25.9)	17.7 (14.7, 20.7)
p-value	<0.001	
Change in 21-item COE scores, item #19		
N	436	178
Baseline, mean (SD)	60.2 (23.2)	58.7 (23.1)
Week 56, mean (SD)	46.0 (23.4)	50.7 (23.6)
Mean change (SD)	-14.3 (27.5)	-7.9 (25.0)
LS mean (SE)	-13.8 (1.2)	-8.5 (1.8)
Diff of LS Mean (95% CI)	-5.3 (-9.2, -1.4)	
p-value	0.007	
Change in IDS-SR Total score		
Baseline, mean (SD)	5.8 (4.8)	6.1 (5.3)
Week 56, mean (SD)	6.0 (5.0)	6.0 (5.0)
Mean change (SD)	0.2 (4.8)	-0.1 (6.4)
LS mean (SE)	0.1 (0.2)	0.0 (0.4)
Diff of LS Mean (95% CI)	0.1 (-0.7, 0.9)	
p-value	0.827	

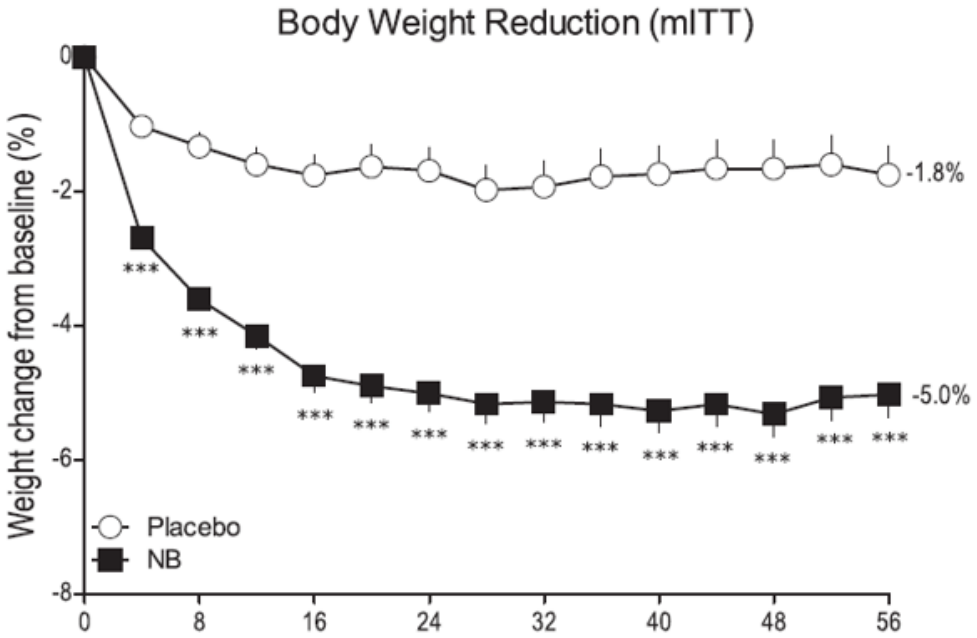
Key: CI, confidence interval; COE, control of eating; HDL, high density lipoprotein; HOMA-IR, homoeostasis model-insulin resistance; hs-CRP, high-sensitivity C-reactive protein; IDS-SR, Inventory of Depressive Symptoms – subject rated; IWQOL-Lite, impact of weight on quality of life-lite; LS, least squares; NB32, naltrexone 32mg plus bupropion; OR, odds ratio; SD, standard deviation; SE, standard error.
Source: Wadden et al. 2010[18]; Orexigen, 2010[63]; EMA, 2014[53]

COR-DM study

Co-primary efficacy analysis: mean percent change in body weight

In the primary population (mITT), patients treated with NB32 lost significantly more weight than placebo-treated patients (LS mean: 5.0% vs 1.8%, respectively; p<0.001) (Figure 13). The difference between groups was significant at Week 4 (the first assessment) and was sustained throughout 56 weeks of treatment (p<0.001 for all visits).

Figure 13: Mean change in body weight, COR-DM study, mITT population



Key: mITT, modified intent-to-treat; NB, naltrexone plus bupropion
Notes: ***, p<0.001; all data are LS Mean (SE)
Source: Hollander et al. 2013[19]

Co-primary efficacy analysis: proportion of patients with ≥5% reduction in body weight

Consistent with percent changes in body weight, more patients treated with NB32 than placebo achieved ≥5% reduction in body weight at Week 56 (44.5% vs 18.9%, respectively; p<0.001) (Table 24) (mITT population).

Table 24: Co-primary efficacy analysis, COR-DM study, mITT population

	NB32 (n=265)	**Placebo (n=159)**
Mean change in body weight		
Baseline weight, mean kg (SD)	106.4 (19.1)	105.0 (17.1)
Week 56 weight, mean kg (SD)	101.0 (19.7)	103.0 (17.3)
Percent change from baseline:		
Mean (SD)	-5.1 (5.7)	-1.8 (4.6)
LS mean (SE)	-5.0 (0.3)	-1.8 (0.4)
Diff of LS mean (95% CI)	-3.3 (-4.3, -2.2)	

p-value	<0.001	
Proportion of patients with ≥5% decrease in body weight		
N (%)	118 (44.5)	30 (18.9)
95% CI	38.5, 50.5	12.8, 25.0
OR (95% CI)	3.4 (2.2, 5.5)	
p-value	<0.001	

Key: CI, confidence interval; diff, difference; LS, least squares; NB32, naltrexone 32mg plus bupropion; SD, standard deviation; SE, standard error
Source: Hollander et al. 2013[19]; Orexigen, 2009[64]; EMA, 2014[53]

As in the previous studies, sensitivity analyses on the co-primary efficacy variables were conducted to address sources of potential bias. Full results of these analyses are presented in Appendix 5. Results of the sensitivity analyses for the LS mean percent change in body weight from baseline to Week 56 were highly consistent with results obtained using the mITT population. In addition, the proportion of patients with ≥5% decrease in body weight were concordant with results for the mITT population.

Secondary efficacy analyses

At 56 weeks, NB32-treated patients had a greater reduction in HbA1c compared to placebo-treated patients (0.6% vs 0.1%, respectively; p<0.001), showing that NB32 can improve glycaemic control in diabetic patients (Table 25). A greater proportion of patients achieved a HbA1c of <7.0% and <6.5% when treated with NB32 compared to placebo (p<0.001 and p=0.004, respectively [Table 25]). Over the course of the study, fewer NB32-treated patients required an increase in dose or the addition of another oral anti-diabetes drug owing to deterioration of glycaemic control (22.3% for NB32-treated patients vs 35.2% for placebo-treated patients; p<0.01). In addition, more NB32-treated patients required a dose reduction in oral anti-diabetes medications and no patients in the NB32 treatment group discontinued the study due to poor glycaemic control, compared to 1.9% of placebo-treated patients.

The proportion of patients with ≥10% decrease in body weight was greater in the NB32 group (18.5%) compared to the placebo group (5.7%).[53] Compared with placebo-treated patients, NB32-treated patients had significantly greater reductions in waist circumference and serum triglyceride concentration and significant increases

in HDL cholesterol. No significant differences were observed between the groups in LDL cholesterol or hs-CRP.

HRQL and PRO measures

At Week 56, the mean change in IWQOL-Lite total score was numerically greater in patients treated with NB32 compared with placebo-treated patients.[64] Although this was not a statistically significant difference, this may reflect the greater burden of disease of this population, compared to patients who are overweight or obese but do not have diabetes. In addition, as it is thought that IWQOL-Lite scores correlate with changes in body weight, this may be due to the slightly smaller magnitude of weight loss in patients with T2DM.This is discussed further in Section 4.13. The reduction in COE questionnaire item #19 score was greater in the NB32 group compared to the placebo group at Week 56 (-11.9 vs -6.9, respectively), indicating improved control of eating after treatment with NB32 even in patients with T2DM.[64]

Table 25: Summary of key secondary endpoints, COR-DM study, mITT population

	NB32 (n=265)	**Placebo (n=159)**
Change in HbA1c (%)		
Baseline, mean (SD)	8.0 (0.8)	8.0 (0.9)
Week 56, mean (SD)	7.3 (1.1)	7.8 (1.2)
LS Mean change (SE)	-0.6 (0.1)	-0.1 (0.1)
p-value	<0.001	
Proportion of patients with HbA1c <7% (%)		
N	222	137
N (%)	98 (44.1)	36 (26.3)
OR (95% CI) [p-value]	2.5 (1.5, 4.0) [<0.001]	
Proportion of patients with HbA1c <6.5% (%)		
N	222	137
N (%)	46 (20.7)	14 (10.2)
OR (95% CI) [p-value]	2.6 (1.4, 5.1) [0.004]	
Proportion of patients with ≥10% decrease in body weight to Week 28		
N (%)	49 (18.5)	9 (5.7)
OR (95% CI) [p-value]	3.8 (1.8, 7.9) [<0.001]	
Change in waist circumference (cm)		
N	208	124

	NB32 (n=265)	Placebo (n=159)
Baseline, mean (SD)	115.6 (12.6)	114.3 (12.4)
Week 56, mean (SD)	110.3 (12.8)	111.3 (12.3)
LS mean (SE)	-5.0 (0.5)	-2.9 (0.6)
p-value	0.006	
Change in fasting insulin levels (µIU/mL)		
N	201	113
Baseline, geometric mean (SD)	15.1 (1.9)	13.8 (1.9)
Week 28, geometric mean	12.9	12.6
LS mean percent change (95% CI)	-13.5 (-19.7, -6.8)	-10.4 (-18.8, -1.1)
p-value	0.563	
Change in fasting HDL cholesterol levels (mg/dL)		
N	222	135
Baseline, mean (SD)	46.2 (10.2)	46.1 (11.5)
Week 56, mean (SD)	49.3 (11.8)	46.0 (12.6)
LS mean (SE)	3.0 (0.5)	-0.3 (0.6)
p-value	<0.001	
Change in fasting LDL cholesterol levels (mg/dL)		
N	220	134
Baseline, mean (SD)	100.2 (34.2)	101.0 (33.9)
Week 56, mean (SD)	99.2 (35.8)	101.0 (37.5)
LS mean (SE)	-1.4 (2.0)	0.0 (2.4)
p-value	0.641	
Change in fasting triglycerides (mg/dL)		
N	222	135
Baseline, geometric mean (SD)	143.3 (1.7)	165.6 (1.6)
Week 56, geometric mean	130.5	158.4
LS mean percent change (95% CI)	-11.2 (-15.6, -6.6)	-0.8 (-7.0, 5.8)
p-value	0.007	
Change in hs-CRP levels (mg/L)		
N	202	119
Baseline, geometric mean (SD)	3.6 (3.0)	3.3 (2.8)
Week 56, geometric mean	2.9	3.0

	NB32 (n=265)	Placebo (n=159)
LS mean percent change (95% CI)	-20.9 (-29.3, -11.5)	-13.3 (-24.9, 0.2)
p-value	0.312	
Change in fasting blood glucose levels (mg/dL)		
N	264	158
Baseline, mean (SD)	160.0 (41.3)	163.9 (44.5)
Week 56, mean (SD)	148.7 (46.7)	158.2 (48.2)
LS mean change (SE)	-11.9 (2.7)	-4.0 (3.4)
p-value	0.065	
Change in HOMA-IR levels		
N	199	112
Baseline, geometric mean (SD)	5.7 (2.0)	5.2 (2.0)
Week 28, geometric mean	4.5 (0.5)	4.5 (0.7)
LS mean percent change (95% CI)	-20.6 (-27.8, -12.6)	-14.7 (-24.7, -3.3)
p-value	0.361	
Change in systolic blood pressure (mmHg)		
Baseline, mean (SD)	125.0 (11.0)	124.5 (9.6)
Week 56, mean (SD)	125.1 (12.7)	123.6 (11.5)
LS mean (SE)	0.0 (0.7)	-1.1 (0.9)
p-value	0.297	
Change in diastolic blood pressure (mmHg)		
Baseline, mean (SD)	77.5 (7.5)	77.4 (7.1)
Week 56, mean (SD)	76.3 (8.7)	75.9 (8.4)
LS mean change (SE)	-1.1 (0.5)	-1.5 (0.6)
p-value	0.582	
Change in IWQOL-Lite total scores		
N	241	153
Baseline, mean (SD)	73.2 (17.2)	73.5 (16.9)
Week 56, mean (SD)	82.5 (15.9)	81.4 (15.4)
Mean change (SD)	9.3 (12.0)	7.9 (11.3)
LS mean (SE)	9.3 (0.7)	7.9 (0.9)
Diff of LS change (95% CI)	1.4 (-0.8, 3.5)	
p-value	0.208	
Change in 21-item COE scores, item #19		

	NB32 (n=265)	Placebo (n=159)
N	225	146
Baseline, mean (SD)	58.0 (22.4)	55.6 (23.5)
Week 56, mean (SD)	45.5 (22.5)	49.5 (21.7)
Mean change (SD)	-12.5 (23.8)	-6.0 (25.0)
LS mean (SE)	-11.9 (1.4)	-6.9 (1.7)
Diff of LS Mean (95% CI)	5.0 (-9.2, -0.7)	
p-value	0.021	
Change in IDS-SR Total score		
N	265	159
Baseline, mean (SD)	8.2 (5.9)	7.8 (5.7)
Week 56, mean (SD)	8.3 (6.6)	6.4 (5.5)
Mean change (SD)	0.1 (6.0)	-1.4 (5.4)
LS mean (SE)	0.0 (0.3)	-1.6 (0.4)
Diff of LS Mean (95% CI)	1.6 (0.6, 2.7)	
p-value	0.002	

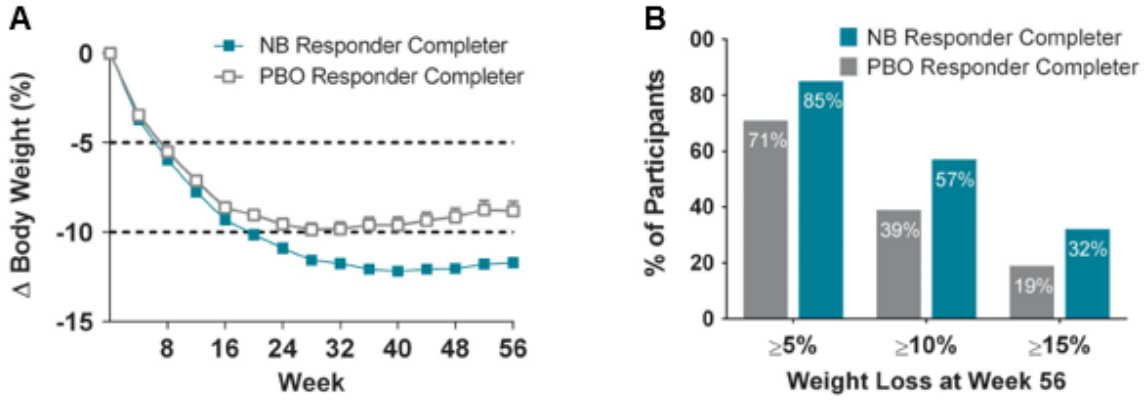
Key: CI, confidence interval; COE, control of eating; FCI, Food Craving Inventory; HbA1c, glycated haemoglobin; HDL, high density lipoprotein; HOMA-IR, homoeostasis model-insulin resistance; hs-CRP, high-sensitivity C-reactive protein; IDS-SR, Inventory of Depressive Symptoms - Subject Rated; IWQOL-Lite, impact of weight on quality of life-lite; LDL, low density lipoprotein; LS, least squares; NB32, naltrexone 32mg plus bupropion; OR, odds ratio; SD, standard deviation; SE, standard error.
Source: Hollander et al. 2013[19]; Orexigen, 2009[64]

Pooled Analysis of COR Trial Responders

A pooled analysis of patient-level data from the four pivotal COR trials was conducted to evaluate early weight loss with NB32 as a predictor of clinically meaningful long-term (Week 56) weight loss of ≥5%.[20] For the evaluation of each early weight loss threshold, responders were defined as participants who lost at least 2, 3, 4 or 5% of baseline weight at Weeks 8, 12 or 16. Non-responders were defined as participants who either gained weight or who did not achieve the responder threshold. Analyses that examined the relationship between participant achievement of early treatment weight loss thresholds and the associated weight loss at Week 56 were conducted on the completers population, defined as patients who had a baseline and Week 56 weight measurement while on study treatment (NB32: n=1,310; placebo: n=763).

Of the patients with observed data at Week 16, 50.8% of those treated with NB32 had lost ≥5% of their baseline body weight, compared with 19.3% of placebo-treated subjects (Week 16 Responders).[20] Week 16 Responders who received NB32 had a high retention rate in the study with 87% completing one year of treatment. At Week 56, the LS mean weight loss (using LOCF methodology) among Week 16 Responder Completers was 11.7%, with 57% of these patients losing ≥10% of their original body weight, as depicted in Figure 14.

Figure 14: Weight loss at Week 16, ≥5% responder completers



Notes: A) weight loss by visit (least-squares mean + standard error); B) proportion of participants with categorical weight loss at Week 56 in the Week 16 ≥5% responder completer population.
Source: Fujioka et al. 2016[20]

The clinically recommended ≥5% weight loss threshold at Week 16 correctly identified 80% (95% CI: 78, 82%) of the participants who would, and would not, achieve ≥5% weight loss at Week 56. The 20% of participants who would not have been correctly identified using this criterion were divided nearly equally among those who would have been inappropriately removed from treatment due to insufficient weight loss at Week 16 (that is, despite ultimately achieving ≥5% at Week 56; n=125) and those who would have inappropriately continued study treatment for 1 year (i.e. participants who achieved the ≥5% at the Week 16 threshold but did not achieve ≥5% weight loss at Week 56; n=132).

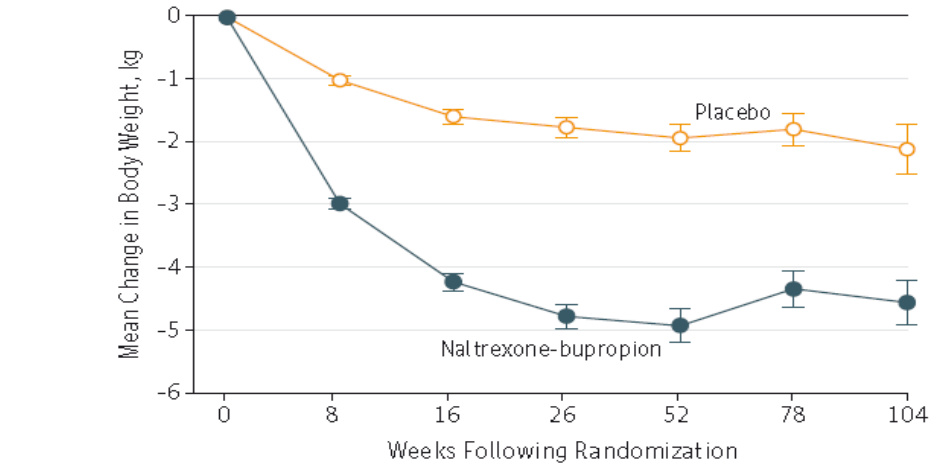
NB-CVOT study

Primary and secondary CV outcomes for the ITT population at the 50% interim analysis, and at the final end-of-study analysis are presented in Appendix 7.

For the 50% interim analysis, the primary prespecified outcome measure, time to first MACE, occurred in 192 patients; 102 (2.3%) in the placebo group and 90 (2.0%) in the NB32 group (HR: 0.88; 99.7% CI: 0.57, 1.34). The components of the primary composite outcome included CV death (0.8% of placebo patients and 0.4% of NB32 patients [HR: 0.50; 99.7% CI: 0.21, 1.19]), nonfatal stroke (0.4% of placebo patients and 0.5% of NB32 patients [HR: 1.10; 99.7% CI: 0.44, 2.78]) and nonfatal myocardial infarction (1.2% in both placebo and NB32 patients [HR: 1.00; 99.7% CI: 0.57, 1.75]). In general, final end-of-study analyses support these data (see Appendix 7).

At trial completion, body weight decreased by a mean of 3.9kg (95% CI: -4.1, -3.7kg) in the NB32 group compared to a mean decrease of 1.2kg (95% CI: -1.3, -1.0kg) in the placebo group, corresponding to reductions of 3.6% and 1.1%, respectively ($p<0.001$ [Figure 15]). The between-group mean difference was 2.7kg (95% CI: −2.9, −2.5kg; $p<0.001$), representing a 2.5% improved reduction (95% CI: −2.8%, −2.3%) in body weight for patients treated with NB32.

Figure 15: Change on body weight during trial



No. of patients							
Placebo	4450	4042	3738	3297	2848	2507	2264
Naltrexone-bupropion	4455	3977	3677	3404	2995	2690	2408

Note: Error bars indicate 95% confidence interval.
Source: Nissen et al. 2016[52]

IGNITE study

The IGNITE study was a Phase IIIb, randomised, open-label, controlled study which assessed the effects of NB32 plus standard management compared to usual care in

adults with obesity.[54] A total of 242 patients were randomised; 153 to NB32 plus standard management and 89 patients to usual care for a total of 26 weeks. NB32 patients not achieving ≥5% weight loss at Week 16 were discontinued, as indicated by product labelling. After Week 26, usual care subjects began treatment with NB32 plus standard management. Assessments continued through Week 78. The primary endpoint was percent change in weight from baseline to Week 26 in the per protocol (PP) population. Other endpoints included percentage of patients achieving ≥5%, ≥10%, and ≥15% weight loss, percent change in weight at Week 78 and AEs necessitating study discontinuation.

Patients assigned to treatment with NB32 plus standard management lost significantly more weight than patients treated with usual care at Week 26 (-9.4% vs -0.94% respectively; p<0.0001). For patients who remained on treatment, the initial weight loss observed at 26 weeks was sustained throughout Week 78, further supporting the maintained effectiveness of NB32 treatment.

4.8 Subgroup analysis

COR-I study

Prospectively defined subgroup analyses were performed on the co-primary efficacy variables, using the mITT population within selected populations by study centre, selected baseline characteristics, BMI, presence of hypertension and dyslipidaemia at baseline.

Overall, the effects of treatment with NB16 and NB32 versus placebo, as measured by the mean percent change in body weight and the proportion of subjects with ≥5% decrease in body weight from baseline to endpoint, were generally consistent across the subgroups defined by demographics (sex, race, and age), study centre, BMI category (<median, ≥median value of 36kg/m^2), and presence of hypertension and dyslipidaemia. Additionally, the magnitude of the treatment effect in fasting triglycerides, HDL cholesterol, and LDL cholesterol levels was similar for NB16- and NB32-treated subjects irrespective of dyslipidaemia diagnosis.

COR-II study

Pre-specified subgroup analyses were performed on the co-primary efficacy variables the results of which showed consistent treatment effects across study centres, demographic variables, BMI category, hypertension and dyslipidaemia.

Subgroup analyses were also performed on selected metabolic parameters. These showed consistent treatment effects from baseline to Week 28 in percent change in fasting triglycerides and mean change in HDL cholesterol and LDL cholesterol, suggesting that the effects of NB32 on lipids are not meaningfully affected by the presence of dyslipidaemia.

COR-BMOD study

Subgroup analyses were performed on the co-primary efficacy variables, and results showed generally consistent treatment effects observed across all demographic variables. This suggests that NB32 was effective regardless of sex, race or age category. In addition, a consistent treatment effect was observed across the study centres, BMI categories, and patients with hypertension and dyslipidaemia.

Post-hoc subgroup analyses were performed on the mean change from baseline to endpoint for fasting triglycerides, fasting HDL cholesterol, and fasting LDL cholesterol using the mITT population within the dyslipidaemia population. Consistent treatment effects from baseline to endpoint in mean fasting triglycerides, HDL cholesterol, and LDL cholesterol levels were observed, suggesting that the effects of NB32 on lipids are not meaningfully affected by the presence of dyslipidaemia.

COR-DM study

Pre-specified subgroup analyses were performed on the co-primary efficacy variables, the results of which showed generally consistent treatment effects across study centre and demographic variables including sex, race, tobacco use, age and age group. In addition, generally consistent treatment effects were observed across BMI categories, HbA1c strata and sulfonylurea pharmacotherapy, suggesting that NB32 compared to placebo was effective regardless of baseline demographics. This was also true for dyslipidaemia and hypertension subgroups.

Consistent treatment effects from baseline to endpoint in geometric mean fasting triglycerides and mean HDL cholesterol and LDL cholesterol were also observed, showing that NB32 was effective regardless of the presence of dyslipidaemia.

No subgroup analyses were conducted in the NB-CVOT or IGNITE studies.

4.9 Meta-analysis

4.9.1 Evidence base

The trial characteristics and analysis specifics for each of the four trials investigating NB32 trials are given in Table 26. The patient populations are broadly similar in three of the trials (COR-I, COR-II and COR-BMOD) where T2DM patients were excluded. The COR-DM trial only included T2DM patients. Out of the three non-T2DM trials, the standard management therapy received in the COR-BMOD trial, was more intensive than in the COR-I and COR-II trials. These differences (the presence or absence of T2DM and the intensity of the diet and exercise programme) between trial designs are likely to explain the heterogeneity in results between the four trials. The trials all consistently used a mITT population, and LOCF rules for patients without data at the analysis time point.

Table 26: NB32 trial characteristics and analysis specifics

Trial	T2DM	Intensive BMOD	Analysis population	Analysis specifics	Sample size (ITT)	Sample size (mITT)
COR-I	Trial excluded T2DM patients	No	mITT – all randomised participants with a baseline weight and one or more post-baseline weight measurement while on study drug	LOCF	NB32: 583 PBO: 581	NB32: 511 PBO: 471
COR-II	Trial excluded T2DM patients	No	mITT – all randomised participants with a baseline weight and one or more post-baseline weight measurement while on study drug	LOCF NB32 patients who had <5% weight loss at visits between Weeks 28–44 were defined as non-responders and were re-randomised to receive NB48 or NB32. Non-responders treated with NB48 were excluded, non-responders treated with NB32 were double weighted.	NB32: 1001 PBO: 495	NB32: 702[a] PBO: 456
COR-BMOD	Trial excluded T2DM patients	Yes	mITT – all randomised participants with a baseline weight and one or more post-baseline weight measurement while on study drug	LOCF	NB32: 591 PBO: 202	NB32: 482 PBO: 193
COR-DM	Trial included only patients with T2DM	No	mITT – all randomised participants with a baseline weight and one or more post-baseline weight measurement while on study drug	LOCF	NB32: 335 PBO: 170	NB32: 265 PBO: 159

Key: BMOD, behaviour modification; COR, Contrave® obesity research; DM, diabetes mellitus; ITT, intent-to-treat; LOCF, last observation carried forward; mITT, modified intent-to-treat; NB32, naltrexone 32mg plus bupropion; NB48, naltrexone 48mg plus bupropion; PBO, placebo; T2DM, Type 2 diabetes mellitus.
Notes: [a], mITT population at 1 year.

The NB-CVOT study was excluded from all meta-analyses, due to the trial design, objective, and patient population, being different from the other studies. As detailed in Section 4.2, the NB-CVOT study was terminated early. Some key differences in trial design are described here. A total of 10,514 patients entered an initial 2-week lead-in period in the NB-CVOT study to identify patients who did not tolerate treatment or exhibited poor compliance. 9,015 patients completed this lead-in period; details of discontinuations within the lead-in period are presented in Table 27.

Table 27: Lead-in period discontinuations in the NB-CVOT study

Reason for discontinuation during the lead-in period	Frequency
Adverse events	543
Did not meet eligibility criteria	425
Withdrew consent	216
Protocol deviation	109
Lost to follow-up	82
Sponsor decision	29
Missing	3
Other reasons	92
Key: CVOT, cardiovascular outcome trial; NB, naltrexone plus bupropion.	

The NB-CVOT study also incorporated a treatment stopping rule in the trial where patients discontinued treatment after 16 weeks if they did not achieve ≥2% reduction in weight, while no stopping rule was used in the COR trials, further details are provided in Section 4.13.[52]

4.9.2 Outcomes and methods

To compare and pool the relative treatment effects between the four trials comparing NB32 and placebo, a frequentist pairwise meta-analysis was performed to assess the following outcomes:

- At least a 5% reduction in weight at 1 year from baseline (the 1-year time point ranged from 52 to 57 weeks). This was a dichotomous outcome.
- Mean % weight change from baseline at 1 year (the 1-year time point ranged from 52 to 57 weeks). This was a continuous outcome.

The frequentist pairwise meta-analysis was performed using R (version 3.3.1) using the metafor package.[65, 66] The pairwise meta-analysis, presents relative treatment effects per trial, and an overall 'pooled' relative treatment effect for placebo vs NB32, calculated using a random effects model.[67] A random effects model was chosen over the fixed effects model as it allows the treatment effect to vary across trials. The random effects model therefore captures the between-trial heterogeneity described in Section 4.9.1. It therefore provided a better estimate of the variation around the overall relative treatment effect estimate compared to a fixed effects model. To further evaluate the trial-heterogeneity, sensitivity analyses were also performed for the three non-T2DM trials, and for the non-T2DM trials excluding the COR-BMOD trial, as patients received intensive behaviour modification. The statistical heterogeneity of the pairwise meta-analysis was assessed using I^2, where the I^2 value describes the percentage of total variation across studies that is due to heterogeneity rather than chance.[68] Table 28 gives an approximate interpretation of between trial heterogeneity based upon I^2 values.

Table 28: Interpretation of I^2

I^2 Value	Between-trial heterogeneity
$0 \leq I^2 < 25$	None to low
$25 \leq I^2 < 50$	Low to moderate
$50 \leq I^2 < 75$	Moderate to high
$75 \leq I^2 \leq 100$	High

Results of the pairwise meta-analysis with I^2 values are presented in forest plots in Section 4.9.3. Forest plots and results for the sensitivity analyses are presented in Appendix 8. The mITT population with LOCF was used for the meta-analysis as described for each trial in Table 27.

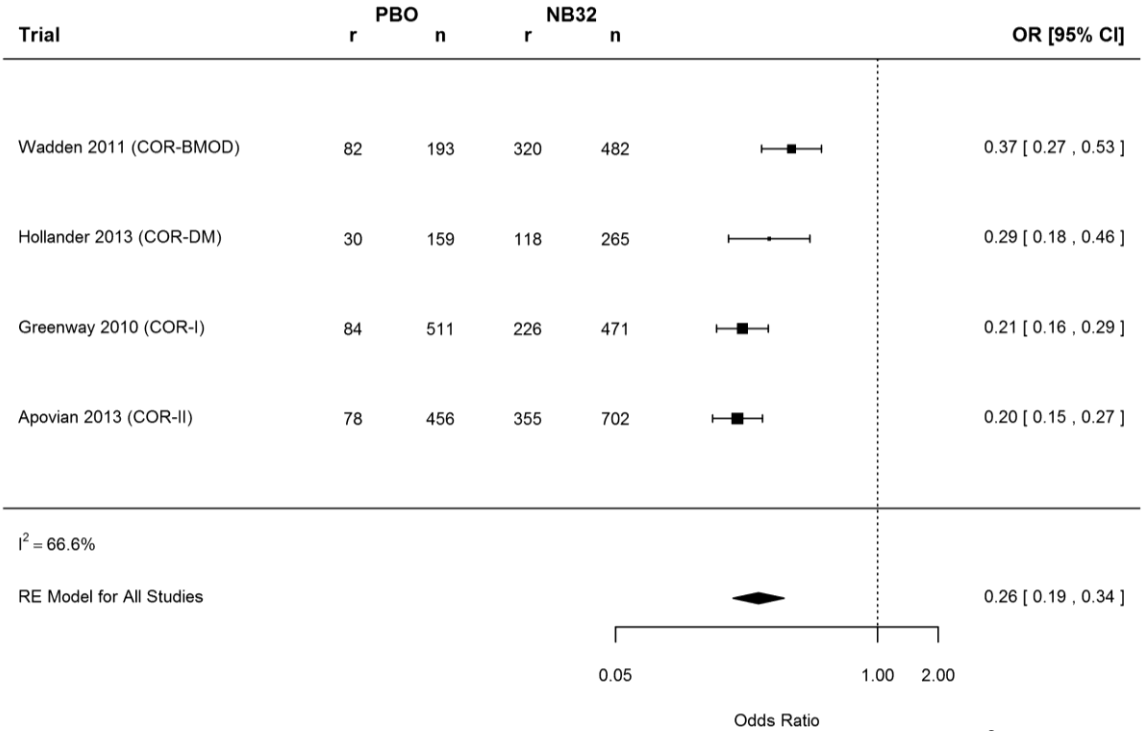
4.9.3 Results

At least 5% reduction in weight

Results are presented as ORs with 95% confidence intervals (CI) on a log scale. An OR less than one favours NB32 over placebo. Figure 16 displays the pairwise meta-analysis results for ≥5% reduction in weight for placebo versus NB32 for all four trials investigating NB32. Across these trials, patients who received placebo had significantly lower odds of achieving a 5% reduction in weight versus NB32 (pooled

OR: 0.26 [95% CI: 0.19, 0.34]). COR-I and COR-II produced similar results, which was expected as the patient populations and treatments in the trials were similar. In the COR-DM and COR-BMOD trials, differences between NB32 and placebo were less pronounced (compared to COR-I and COR-II). Nevertheless, results were still significantly in favour of NB32. For COR-BMOD in particular, the higher OR reflects that more placebo patients lost at least 5% of their initial weight due to the more intensive behaviour modification program relative to placebo patients in the other studies. The I^2 value indicates moderate-high heterogeneity, which is likely to be due to the differences observed in the COR-BMOD and COR-DM trials compared to the COR-I and COR-II trials for placebo versus NB32.

Figure 16: Forest plot of ≥5% reduction in weight for placebo versus NB32 (all trials)



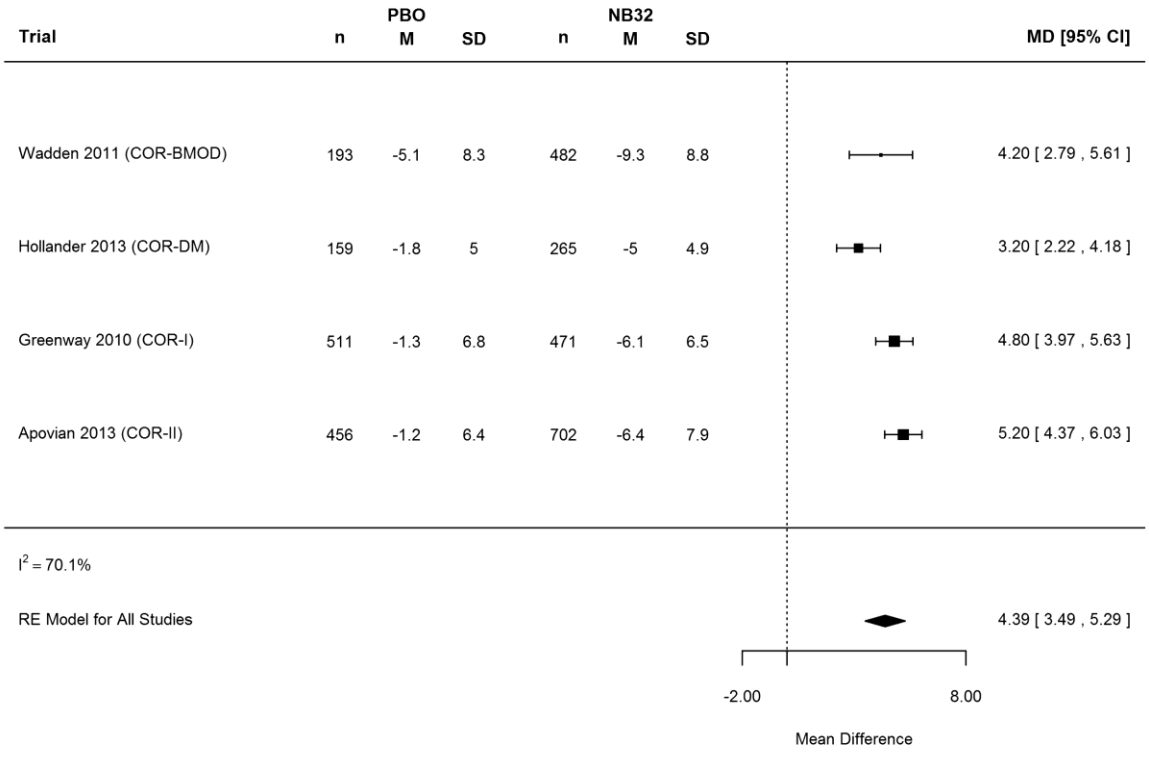
Key: BMOD, intensive behaviour modification; CI, confidence interval; COR, Contrave® obesity research; DM, diabetes mellitus; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; r, number of patients achieving ≥5% reduction in weight; RE, random effects.

Percentage weight change from baseline

Results are presented as mean differences (MDs) with 95% CIs on a linear scale. The mean differences are calculated as mean % weight change from baseline at 1 year in the placebo group minus the mean % weight change from baseline at 1 year in the NB32 group. A MD greater than 0 favours NB32 over placebo. Figure 17

displays the pairwise meta-analysis results for % weight change from baseline for placebo versus NB32. For all four trials, patients who received placebo had a significantly smaller % reduction in weight (at 1 year compared to baseline) versus NB32 (pooled MD: 4.39 [95% CI: 3.49, 5.29]). COR-I and COR-II produced similar results, which was expected as the patient populations and treatments in the trials were similar. The COR-DM trial produced lower mean differences of response compared to the COR-I and COR-II trials for placebo versus NB32. The MD in the COR-BMOD trial is also lower than the COR-I and COR-II trials for placebo versus NB32; however, there is more uncertainty around this estimate. The I^2 value indicates moderate-high heterogeneity, which is likely to be due to the lower MD observed in the COR-DM trials compared to the COR-I and COR-II trials for placebo versus NB32.

Figure 17: Forest plot for % weight CFB for placebo versus NB32 (all trials)



Key: BMOD, behaviour modification; CFB; change from baseline; COR, Contrave® obesity research; DM, diabetes mellitus; M, mean; MD, mean difference; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; SD, standard deviation; RE, random effects.

4.10 Indirect and mixed treatment comparisons

Orlistat is currently the only orally effective pharmaceutical product for weight management available through NHS England, therefore the relative efficacy between orlistat and NB32 is of interest. However, in the absence of head-to-head trials between orlistat and NB32, an indirect treatment comparison (ITC) using placebo as a common comparator was required. A SLR was performed to identify RCTs for both NB32 and orlistat to be used within the ITC.

4.10.1 Search strategy

The search strategy used to identify RCT evidence for NB32 and orlistat 120mg TID is described in Section 4.1.

4.10.2 Study selection

Full eligibility criteria applied to the systematic search results identifying the clinical evidence base of randomised trials are outlined in Section 4.1.

In addition to the criteria listed for trial inclusion/exclusion in the SLR (Section 4.1, Table 10), eligibility criteria were applied to confirm a final set of trials suitable for the network of evidence for the ITC. The additional criteria excluded trials and treatment arms for the following reasons:

- Treatment arm is not of interest
- Treatment group is not administered at recommended dosage
- Trial reduces to single treatment arm once other arms are pooled or excluded
- Trial reports no relevant outcome data
- Trial excludes patients during a lead-in period due to weight loss criteria or treatment compliance
- Trial has a wait list control group as a comparator arm in which patients receive no pharmaceutical treatment or standard management

For the performed analyses, NB and orlistat were evaluated at their recommended doses detailed in the summary of product characteristics for each treatment[10, 29]:

- NB – naltrexone 32mg/day prolonged release plus bupropion 360mg/day prolonged release (NB32)
- Orlistat – 120mg three times a day (TID)

ITC were performed to compare NB32 and orlistat for the following outcomes:

- Mean % weight change from baseline at 1 year (the 1-year time point ranged from 52 to 57 weeks [continuous outcome])
- At least 5% reduction in weight at 1 year from baseline (the 1-year time point ranged from 52 to 57 weeks [dichotomous outcome])

Table 29, shows the list of trials, along with treatments and available outcome data that were included in the analyses; individual treatment arms that were excluded from these studies are detailed in Table 30. Appendix 9 details a list of 13 trials identified in the SLR that were not considered part of the analyses. The maximum evidence base for each outcome, following the additional exclusion is given in the network of evidence presented in Figure 18.

Table 29: Evidence base: trials, treatments and outcomes

Trial (NT=20)	Trial duration	Treatment		Analysis population					Outcome	
		Arm 1	Arm 2	Trials without a lead-in period	Trial excludes T2DM patients[a]	T2DM is part of trial inclusion criteria	Trials without a high proportion of comorbidities[b]	Standard management without intensive BMOD[c]	≥5% reduction in weight	Mean % weight CFB
Apovian 2013[17] (COR-II; NCT00567255)[d]	56 weeks	PBO	NB32	✓	✓	-	✓	✓	✓	✓
Greenway 2010[16] (COR-I; NCT00532779)	56 weeks	PBO	NB32	✓	✓	-	✓	✓	✓	✓
Hollander 2013[19] (COR-DM; NCT00474630)	56 weeks	PBO	NB32	✓	-	✓	-	✓	✓	✓
Wadden 2011[18] (COR-BMOD; NCT00456521)	56 weeks	PBO	NB32	✓	✓	-	✓	-	✓	✓
Astrup 2012[69] (NN8022-1807 study group; NCT00422058 [extension study: NCT00480909)	54 weeks (2-week lead-in period and 52-week treatment phase [weeks 20–52 were part of an extension study])	PBO	ORL 120mg TID	-	✓	-	✓	✓	✓	✓
Bakris 2002[70]	52 weeks	PBO	ORL 120mg TID	✓	-	-	-	✓	✓	✓

Trial (NT=20)	Trial duration	Treatment		Analysis population					Outcome	
		Arm 1	Arm 2	Trials without a lead-in period	Trial excludes T2DM patients[a]	T2DM is part of trial inclusion criteria	Trials without a high proportion of comorbidities[b]	Standard management without intensive BMOD[c]	≥5% reduction in weight	Mean % weight CFB
Berne 2005[71] (OST2D study group)	54 weeks (2-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	✓	-	✓	✓	✓
Broom 2002[72] (UKM study group)	54 weeks (2-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	✓	✓	✓	✓
Derosa 2003[73]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO[e]	ORL 120mg TID[f]	-	-	-	-	✓	-	✓
Derosa 2010[74]	52 weeks	PBO	ORL 120mg TID	✓	-	✓	-	✓	-	✓
Gotfredsen 2001[75] (EM Study-I)	52 weeks (4-week lead-in period and 48-week treatment period)	PBO	ORL 120mg TID	-	-	-	✓	✓	-	✓
Karhunen 2000[76] (EM Study-II)	108 weeks (4-week lead-in period and two 52-week treatment periods)	PBO	ORL 120mg TID	-	✓	-	✓	✓	-	✓
Kelley 2002[77]	54 weeks (2-week screening and 52-week treatment phase)	PBO	ORL 120mg TID	✓	-	✓	-	✓	✓	✓

Trial (NT=20)	Trial duration	Treatment		Analysis population					Outcome	
		Arm 1	Arm 2	Trials without a lead-in period	Trial excludes T2DM patients[a]	T2DM is part of trial inclusion criteria	Trials without a high proportion of comorbidities[b]	Standard management without intensive BMOD[c]	≥5% reduction in weight	Mean % weight CFB
Lindgarde 2000[78]	54 weeks (2-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	-	✓	✓	✓
Lucas 2003[79]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	-	✓	-	✓
Mathus-Vliegen 2006[80]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	✓	-	✓	✓	-	✓
Miles 2002[51]	54 weeks (2-week screening period and 52-week treatment phase)	PBO	ORL 120mg TID	✓	-	✓	-	✓	✓	✓
Reaven 2001[81]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO[g]	ORL 120mg TID[g]	-	✓	-	✓	✓	-	✓
Swinburn 2005[82]	56 weeks (4-week lead-in period plus 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	✓	✓	-	✓
Torgerson 2004[83] (XENDOS)	208 weeks	PBO	ORL 120mg TID	✓	✓	-	✓	-	✓	✓

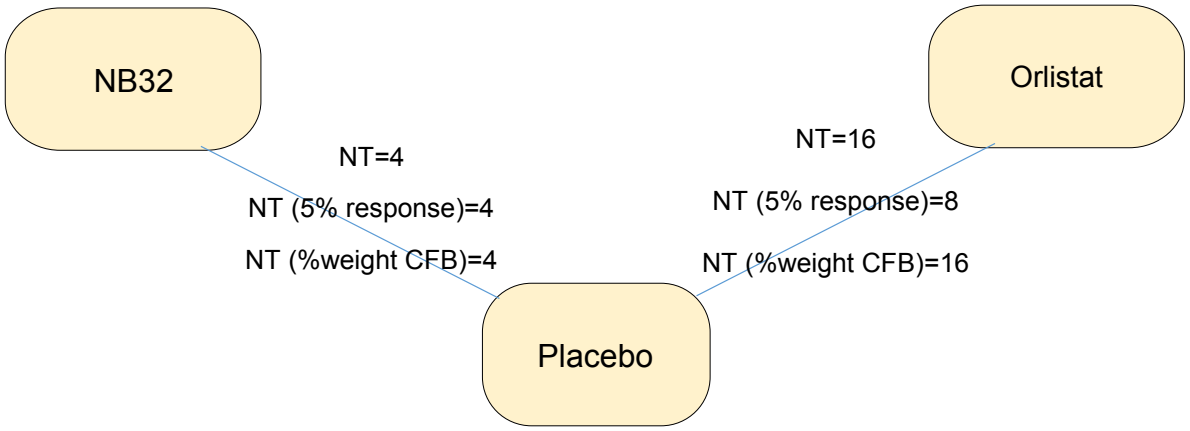
Trial (NT=20)	Trial duration	Treatment		Analysis population					Outcome	
		Arm 1	Arm 2	Trials without a lead-in period	Trial excludes T2DM patients[a]	T2DM is part of trial inclusion criteria	Trials without a high proportion of comorbidities[b]	Standard management without intensive BMOD[c]	≥5% reduction in weight	Mean % weight CFB
Total NB32 trials				**4**	**3**	**1**	**3**	**3**	**4**	**4**
Total ORL trials				**5**	**5**	**4**	**8**	**15**	**8**	**16**
Total trials				**9**	**8**	**5**	**11**	**18**	**12**	**20**

Key: BMOD, behaviour modification; CFB, change from baseline; COR; Contrave® obesity research; DM, diabetes mellitus; EM, European multicentre; FV, fluvastatin; NB, naltrexone plus bupropion; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion; NB48, naltrexone 48mg plus bupropion; NT, number of trials; ORL, orlistat 120mg TID; PBO, placebo; SM, Swedish Multimorbidity; T2DM, Type 2 diabetes mellitus; TID, three times a day. UKM, UK Mulitmorbidity; XENDOS, Xenical in the prevention of diabetes in obese subjects.
Notes: [a], As per the trial exclusion criteria; [b], High proportion of comorbidities were defined as in Section 4.10.3, [c], Intensive BMOD defined as in Section 4.10.4; [d], Non-responders in the Apovian 2013 trial were re-randomised to either NB32 or NB48. Non-responders who received NB48 after 32 weeks were not included in the analysis, and patients who received NB32 were double weighted in the analysis; [e], PBO and PBO+FV have been pooled together; [f], ORL 120mg TID and ORL120mg TID+FV have been pooled together; [g], Trial presents arm data split by whether patients had syndrome X, and patients with/without syndrome X were pooled for each treatment.

Table 30: Treatment arms excluded from analysis

Trial	Treatment arm	Reason for exclusion
Greenway 2010 (COR-I; NCT00532779)	NB16	NB16 below recommended dosage (NB32)
Astrup 2012 (NN8022-1807 study group)	LIRA 1.2/1.8/2.4/3.0mg/day	LIRA not treatment of interest
Key: COR, Contrave® obesity research; LIRA, liraglutide; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion.		

Figure 18: Network of evidence



Key: CFB, change from baseline; NB32, naltrexone 32mg plus bupropion; NT, number of trials.
Notes: 5% response defined as ≥5% reduction in weight from baseline at 1 year.

4.10.3 Methods, outcomes, and data of included studies

At least 5% reduction in weight at 1 year is a dichotomous outcome. ORs were therefore used as the outcome measure. ORs are presented on a log scale as relative treatment effects are log transformed prior to analysis. The log odds ratio (LOR) of the treatment effects are approximately normally distributed and therefore presenting the plots on a log scale produces near symmetrical confidence intervals. The data synthesised for the analysis of ≥5% reduction in weight at 1 year is presented in Table 31.

Table 31: Data synthesised in analyses for ≥5% reduction in weight at 1 year

Study name (trial name)	Arm 1	Arm 2	n1	r1	n2	r2
Apovian 2013 (COR-II)	NB32	PBO	702	355	456	78
Greenway 2010 (COR-I)	NB32	PBO	471	226	511	84

Study name (trial name)	Arm 1	Arm 2	n1	r1	n2	r2
Hollander 2013 (COR-DM)	NB32	PBO	265	118	159	30
Wadden 2011 (COR-BMOD)	NB32	PBO	482	320	193	82
Astrup 2012 (NN8022-1807 study group)	ORL	PBO	95	42	98	27
Bakris 2002	ORL	PBO	267	122	265	60
Berne 2005 (OST2D study group)	ORL	PBO	111	51	109	12
Broom 2002 (UKM study group)	ORL	PBO	259	144	263	64
Derosa 2003	ORL	PBO	NR			
Derosa 2010	ORL	PBO	NR			
Gotfredsen 2001 (EM Study-I)	ORL	PBO	NR			
Karhunen 2000 (EM Study-II)	ORL	PBO	NR			
Kelley 2002	ORL	PBO	266	87	269	35
Lindgarde 2000 (SM Study)	ORL	PBO	190	103	186	76
Lucas 2003	ORL	PBO	NR			
Mathus-Vliegen 2006	ORL	PBO	NR			
Miles 2002	ORL	PBO	250	98	254	40
Reaven 2001	ORL	PBO	NR			
Swinburn 2005	ORL	PBO	NR			
Torgerson 2004 (XENDOS)	ORL	PBO	1640	1194	1637	738

Key: BMOD, behaviour modification; COR, Contrave® obesity research; DM, diabetes mellitus; EM, European multicentre; n, number of patients; NB32, naltrexone 32mg plus bupropion; NR, not reported; OST2D, Orlistat Swedish Type 2 diabetes; ORL, orlistat 120mg; r, number of patients achieving ≥5% reduction in weight; TID; three times a day; PBO, placebo; SM, Swedish Multimorbidity; UKM, UK Multimorbidity; XENDOS, Xenical® in the prevention of diabetes in obese subjects.

Mean % weight change from baseline at 1 year is a continuous outcome. MDs were used as the outcome measure for the analysis of mean % weight change from baseline and are presented on a linear scale as it is assumed that % weight change from baseline is normally distributed. The data synthesised for the analysis of % weight change from baseline at 1 year is presented in Table 32. Some data imputations were required to maximise inclusion of evidence in the analyses, and the methods of imputation are described in Appendix 10.

Table 32: Data synthesised in analysis for % weight CFB-1 year

Study name (trial name)	Arm 1	Arm 2	n1	M1	SE1	n2	M2	SE2
Apovian 2013 (COR-II)	NB32	PBO	702	-6.40	0.30	456	-1.20	0.30
Greenway 2010 (COR-I)	NB32	PBO	471	-6.10	0.30	511	-1.30	0.30
Hollander 2013 (COR-DM)	NB32	PBO	265	-5.00	0.30	159	-1.80	0.40
Wadden 2011 (COR-BMOD)	NB32	PBO	482	-9.30	0.40	193	-5.10	0.60
Astrup 2012 (NN8022-1807 study group)	ORL	PBO	95	-4.06[a]	0.70[b]	98	-2.06[a]	0.68[b]
Bakris 2002	ORL	PBO	267	-5.34[a]	0.39	265	-2.66[a]	0.39
Berne 2005 (OST2D study group)	ORL	PBO	111	-5.00	0.64[b]	109	-1.80	0.65[b]
Broom 2002 (UKM study group)	ORL	PBO	259	-5.80	0.48	263	-2.30	0.38
Derosa 2003[c]	ORL	PBO	49	-10.47	0.14	47	-8.80	0.13
Derosa 2010	ORL	PBO	113	-10.05[d]	0.79[e]	121	-2.84[d]	0.75[e]
Gotfredsen 2001 (EM Study-I)	ORL	PBO	16	-9.57[d]	4.60[e]	14	-8.15[d]	3.17[e]
Karhunen 2000 (EM Study-II)	ORL	PBO	36	-13.35[d]	2.09[e]	36	-8.84[d]	2.68[e]
Kelley 2002	ORL	PBO	266	-3.76	0.26	269	-1.22	0.30
Lindgarde 2000 (SM Study)	ORL	PBO	190	-5.90	0.40	186	-4.60	0.40
Lucas 2003	ORL	PBO	256	-10.04[a]	0.40	188	-6.15[a]	0.50
Mathus-Vliegen 2006	ORL	PBO	10	-9.90	1.35	9	-9.90	2.76
Miles 2002	ORL	PBO	250	-4.60	0.30	254	-1.70	0.20
Reaven 2001[f]	ORL	PBO	156	-8.96	0.63	91	-6.95	0.67
Swinburn 2005	ORL	PBO	170	-4.55[a]	0.59	169	-0.84[a]	0.32
Torgerson 2004 (XENDOS)	ORL	PBO	1640	-9.60[a]	0.17[b]	1637	-5.61[a]	0.17[b]

Study name (trial name)	Arm 1	Arm 2	n1	M1	SE1	n2	M2	SE2
Key: BMOD, behaviour modification; CFB, change from baseline; COR, Contrave® obesity research; DM, diabetes mellitus; EM, European multicentre; FV, fluvastatin; M, mean; n, number of patients; NB32, naltrexone 32mg plus bupropion; NR, not reported; OST2D, Orlistat Swedish Type 2 diabetes; ORL, orlistat 120mg TID; PBO, placebo; SD, standard deviation; SE, standard error; SM, Swedish Multimorbidity; UKM, UK Multimorbidity; XENDOS, Xenical® in the prevention of diabetes in obese subjects; **Notes:** [a], Estimated from mean baseline weight and mean weight CFB; [b], Estimated from pooled 'average' of other treatment arm SD's; [c], ORL 120mg TID and ORL120mg TID+FV have been pooled, and PBO and FV arms have been pooled; [d], Estimated from mean baseline weight and mean weight at 12 months; [e], Estimated from mean baseline weight SD and mean weight at 12 months SD; [f], Treatment arm data were pooled for syndrome X and non-syndrome X patients.								

Baseline characteristics for each of the trials and a summary of the baseline characteristics by treatment group are presented in Appendix 11.

Populations of included trials

The licence agreements for both NB32 and orlistat both specify that if a patient is overweight and not obese they must also have at least one comorbidity. The search strategy detailed in Section 4.1 detailed that patients in all included trials must be either overweight with comorbidities, or obese with or without comorbidities.

For both NB32 and orlistat the licence agreements specify treatment stopping rules based upon patient's response to treatment. The licence agreements specify that patients should be advised to stop treatment if they did not achieve at least 5% reduction in weight after receiving orlistat for 12 weeks, or NB32 for 16 weeks (NB32 has a 4-week escalation period prior to patients receiving the full dose). In the case of NB32, this discontinuation rule was based on post-hoc analyses of the four pivotal trials and could therefore not be implemented in the trials themselves. The orlistat trials also did not include a discontinuation rule, as stated in the license. As such, this is a key difference between the trials and the licensed agreement; further details are provided in Section 4.13. The treatment and evaluations of the patients within the trials therefore do not match the practice of the licenced treatments. This limitation is true for both NB32 and orlistat. Therefore, for the purposes of ITC, we made the assumption that despite the trials not applying the stopping rules given in the licence agreement, the relative treatment effects remain unaffected, and that the direction of any small bias is unknown.

Heterogeneity in patient populations

Many patients in the included trials had at least one comorbidity, most notably hypertension, dyslipidaemia or T2DM. T2DM is of particular interest, as it has been observed that weight loss in patients with T2DM may occur at a slower rate[84], and that some intensive therapy for the treatment of diabetes with certain medications may result in weight gain.[85] To investigate the effect of T2DM and to populate the economic model (in which results from the ITC are applied according to individual patient T2DM status), all the analyses and sensitivity analyses were therefore performed separately for (where data were available):

- Trials where T2DM is part of the trial inclusion criteria (T2DM analysis)
- Trials where T2DM is part of the trial exclusion criteria (non-T2DM analysis)
- All trials regardless of T2DM (any T2DM analysis)

To assess the effects of weight loss in trials where a large proportion of patients had comorbidities, a sensitivity analysis was performed excluding trials where ≥75% of patients had at least one comorbidity (hypertension, dyslipidaemia, or T2DM; sensitivity analysis [SA] number 1 [SA1]).

4.10.4 Risk of bias

The quality assessment of the four COR trials are presented in Appendix 4. Quality assessment of the orlistat RCT's included in the network meta-analysis (NMA) are presented in Appendix 12.

Lead-in periods

Of the 20 studies included in the analyses, 11 of the trials investigating orlistat enrolled patients into a lead-in period prior to randomisation in which no patients were excluded due to lack of efficacy or treatment compliance. All four NB32 trials did not have a lead-in period. The lead-in period is a period in which patients received some form of non-active therapy to start weight loss and assess treatment compliance and tolerance. It is therefore possible that a proportion of weight loss may have occurred prior to receipt of the randomised treatment. There is heterogeneity between trials with respect to the duration and the therapies received during the lead-in periods. In some cases, it is also unclear whether weight change from baseline was recorded from the start of randomisation or the start of the lead-in

period. As it is unclear what the effect of lead-in periods would be on results, sensitivity analyses were performed where trials incorporating lead-in periods were excluded (SA2).

Behaviour modification therapy

In each of the trials patients received NB32 as an adjunct to standard management, consisting of diet instruction, advice on behaviour modification and physical activity suggestions (further details are provided in Section 4.3. As in clinical practice, the specific type and intensity of such standard management varied between the trials, although treatment arms within the same trial received the same standard management. For the analysis, it was therefore assumed that the additional treatment benefit from the standard management was additive but that the relative treatment effect between treatment arms would be unaffected. Given the differences between the standard management received, the effect of the intensive, behaviour modification received in the COR-BMOD trial was investigated. There were few pre-existing criteria regarding the definition of 'intensive' behaviour modification within standard management; therefore, the separation of studies by behaviour modification intensity was somewhat subjective. Ara et al. utilised the following criteria in consideration of behaviour modification intensity[14]:

- Standard – patients had one visit with general dietary/exercise advice given or patients given a lifestyle leaflet
- Enhanced – more than just one visit with more than just advice.

This definition, would have resulted in most studies for both orlistat and NB32 being considered 'intensive'. Therefore, we considered less strict criteria to isolate those studies that may be considered drastically different to the standard management therapy received in most studies. Criteria used to elicit these studies were multi-disciplinary; based on the number of follow-up appointments with a medical/dietary professional; detail and severity regarding the prescription of dietary recommendations; and the level of physical activity participants were encouraged to follow. The following studies were identified as considering 'intensive' behaviour modification.

- COR-BMOD the additional therapy received is defined in Section 4.3

- XENDOS[83] – patients received dietary counselling every two weeks for the first 6 months of treatment, and monthly visits thereafter; in addition to an 800kcal deficit diet plan with a physical activity target to walk at least 1 extra kilometre a day.

For these reasons the standard management in both studies were considered to comprise of ‘intensive’ behaviour modification. However, it is acknowledged that the exclusion of studies due to the intensity of standard management is subjective; therefore, sensitivity analyses were performed where studies with ‘intensive’ behaviour modification were excluded (SA3). To evaluate the effects of intensive behaviour modification without trials with lead-in periods, sensitivity analyses were performed where trials with lead-in periods or ‘intensive’ behaviour modification were excluded (SA4).

4.10.5 Methods of analysis and presentation of results

Summary of performed analyses

Table 33 details the list of performed analyses for all outcomes with respect to T2DM. Table 34 and Table 35 detail the number of trials that report data for each outcome by analysis.

Table 33: Performed analyses

Analysis	Trials with patients with T2DM only		Trials excluding patients with T2DM		All trials regardless of T2DM	
	Bayesian NMA	Frequentist pairwise meta-analysis	Bayesian NMA	Frequentist pairwise meta-analysis	Bayesian NMA	Frequentist pairwise meta-analysis
Base case: All trials included	✓	✓	✓	✓	✓	✓
SA1: Trials with 'high' comorbidities were excluded	-[a]	✓	-[b]	-[b]	✓	✓
SA2: Trials with lead-in periods were excluded	✓	✓	✓	✓	✓	✓
SA3: Trials with intensive BMOD were excluded	-[b]	✓	✓	✓	✓	✓
SA4: Trials with lead-in periods or intensive BMOD were excluded	-[c]	✓	-[a]	✓[d]	✓	✓

Key: BMOD, behaviour modification; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.

Notes: [a], Insufficient data available to perform analysis; [b], Analysis not performed as evidence base the same as the base case analysis; [c], Analysis not performed as evidence base the same as SA2; [d], Analysis only performed for NB32.

Table 34: Number of studies reporting data for ≥5% reduction in weight at 1 year

Analysis	Trials with patients with T2DM only		Trials excluding patients with T2DM		All trials regardless of T2DM	
	NB32	ORL	NB32	ORL	NB32	ORL
Base case: All trials included	1	3	3	2	4	8
SA1: Trials with 'high' comorbidities were excluded	0[a]	0[a]	3[b]	2[b]	3	3
SA2: Trials with lead-in periods were excluded	1	2	3	1	4	4
SA3: Trials with intensive BMOD were excluded	1[b]	3[b]	2	1	3	7
SA4: Trials with lead-in periods or intensive BMOD were excluded	1[c]	2[c]	2[a]	0[a]	3	3

Key: BMOD, behaviour modification; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; SA, sensitivity analysis T2DM, Type 2 diabetes mellitus.
Notes: [a], Insufficient data available to perform analysis; [b], Analysis not performed as evidence base the same as the base case analysis; [c], Analysis not performed as evidence base the same as SA2.

Table 35: Number of studies reporting data for mean % weight CFB at 1 year

Analysis	Trials with patients with T2DM only		Trials excluding patients with T2DM		All trials regardless of T2DM	
	NB32	ORL	NB32	ORL	NB32	ORL
Base case: All trials included	1	4	3	5	4	16
SA1: Trials with 'high' comorbidities were excluded	0[a]	0[a]	3[b]	5[b]	3	8
SA2: Trials with lead-in periods were excluded	1	3	3	1	4	5
SA3: Trials with intensive BMOD were excluded	1[b]	4[b]	2	4	3	15
SA4: Trials with lead-in periods or intensive BMOD were excluded	1[c]	3[c]	2[a]	0[a]	3	4

Key: BMOD, behaviour modification; CFB, change from baseline; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; SA, sensitivity analysis T2DM, Type 2 diabetes mellitus.
Notes: [a], Insufficient data available to perform analysis; [b], Analysis not performed as evidence base the same as the base case analysis; [c], Analysis not performed as evidence base the same as SA2.

Analysis methods

Frequentist pairwise meta-analyses

As a first step to compare the relative treatment effects between trials comparing the same treatments, pairwise meta-analyses were performed for both outcomes and each analysis using the methods described in Section 4.9.2. The results of the frequentist pairwise meta-analyses are presented as forest plots in the Appendix 13.

Bayesian network meta-analyses

A Bayesian NMA was performed for each outcome synthesising data from the identified database (detailed in Table 31 and Table 32). Bayesian analyses rely on Markov Chain Monte Carlo (MCMC) methods, combining prior distributions with the data to construct a posterior distribution of parameters of interest upon which to base summary results.

All models were fitted using the freely available software WinBUGS (version 14)[86], using R (version 3.3.1) as an interface to create relevant output.[65] An initial 50,000

iterations were discarded as the 'burn-in' period, which was assessed by running two chains using different starting values and assessing convergence using Brooks-Gelman-Rubin plots.[87] Then, 10,000 convergence diagnosis and output analysis (CODA) samples (posterior distribution) were retained upon which to base summary estimates. In total, 10,000 samples were deemed sufficient for each of the different analyses as the Monte Carlo error was less than 5% of the standard deviation.[88] Therefore, the samples could be used directly in the economic model, preserving the correlation between treatment effects and avoiding the need to make assumptions regarding the shape of the posterior distribution.

Autocorrelation was assessed to determine whether samples were highly correlated, a thinning interval of 5 was applied to ensure that the chain was mixing well and was representative of the posterior distribution. The goodness-of-fit was assessed using the total residual deviance; a chi square test was used to test whether the number of data points in each model was significantly lower than the total residual deviance.

Random effects were considered for all analyses, and may be preferred to fixed effects, where appropriate; however, random effect results are only presented for the any T2DM analysis. Random effect results are not presented for the T2DM only and non-T2DM analyses, as the models failed to update effectively using the recommended priors, likely due to the low number of studies.

The treatment effect model, estimates the relative efficacy between the three treatments included in the evidence base. Different models are used for each of the two outcomes. The two models detailed below are from the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 2.[89] The WinBUGS code for the models used is provided in Appendix 14.

Treatment effects model – ≥5% reduction in weight

As ≥5% reduction in weight at 1 year is a dichotomous outcome, a binomial likelihood was fitted to the data.

$$r_{ik} \sim Binomial(n_{ik}, p_{ik}),$$

where r_{ik} is the number of patients achieving ≥5% reduction in weight, out of a total sample size n_{ik}, and p_{ik} is the probability of an event occurring in arm k of trial i. A logit link function maps these probabilities (bounded between 0 and 1) into a

continuous measure (bounded between minus and plus infinity). The probability of ≥5% reduction in weight p_{ik} on the logit scale was modelled as:

$$log\ odds = \ logit(p_{ik}) = \ln\left(\frac{p_{ik}}{1-p_{ik}}\right) = \mu_i + \delta_{i,bk} I_{\{k\neq 1\}}$$

for fixed effects analysis and

$$log\ odds = \ logit(p_{ik}) = \ln\left(\frac{p_{ik}}{1-p_{ik}}\right) = \mu_i + d_{i,bk} I_{\{k\neq 1\}}$$

for random effects analysis, where:

$$I_{\{u\}} = \begin{cases} 1\ if\ u\ is\ true \\ 0\ otherwise \end{cases}$$

Trial-specific baseline effects (log-odds of the outcome of the trial control arm) are defined as μ_i.The trial-specific treatment effects (LORs) of arm k relative to arm b in the trial are defined as $d_{i,bk}$. Where random effects have been fitted, the trial specific LORs arise from a common distribution:

$$\delta_{i,bk} \sim N(d_{bk}, \sigma^2),$$

where σ^2 is the between trial variance, and d_{bk} is the estimated mean treatment effect of arm k relative to arm b. For binomial models for the between trial deviation, σ, values of 0 to 0.5 are reasonable and represent mild heterogeneity, values of 0.5 to 1 represent fairly high heterogeneity, and values greater than 1 represent fairly extreme heterogeneity.[90] Table 36 gives the prior distributions used for the analysis of ≥5% reduction in weight at 1 year.

Table 36: Prior distribution used for analysis of ≥5% reduction in weight – 1 year

Model	Between-trial deviation (σ)	Treatment effects (d)	Trial baseline effects (μ)
Fixed effects	NA	$N(0,10000)$	$N(0,10000)$
Random effects	$U(0,2)$	$N(0,10000)$	$N(0,10000)$
Key: *N*, normal distribution; NA, not applicable; *U*, uniform distribution.			

Treatment effect model – mean % weight change from baseline

Mean % weight change from baseline at 1 year is a continuous outcome, therefore a normal likelihood was fitted to the data:

$$y_{ik} \sim N(\theta_{ik}, se_{ik}^2),$$

where y_{ik} is the sample mean % weight change from baseline with standard error se_{ik}, and θ_{ik} is the estimated mean in arm k of trial i. The mean % weight change from baseline θ_{ik} on a natural scale (identity link function used) was modelled as:

$$\theta_{ik} = \mu_i + d_{i,bk} I_{\{k \neq 1\}}$$

for fixed effect analysis and

$$\theta_{ik} = \mu_i + \delta_{i,bk} I_{\{k \neq 1\}}$$

for random effects analysis, where:

$$I_{\{u\}} = \begin{cases} 1 \; if \; u \; is \; true \\ 0 \; otherwise \end{cases}$$

Trial-specific baseline effects (mean % weight change from baseline of the trial control arm) are defined as μ_i. The trial-specific treatment effects (MDs) of arm k relative to arm b in the trial are defined as $d_{i,bk}$. Where random effects have been fitted, the trial specific MDs arise from a common distribution:

$$\delta_{i,bk} \sim N(d_{bk}, \sigma^2),$$

where σ^2 is the between trial variance, and d_{bk} is the estimated mean treatment effects of arm k relative to arm b. Table 37 gives the prior distributions used for the analysis of mean % weight change from baseline.

Table 37: Prior distribution used for analysis of mean % weight CFB

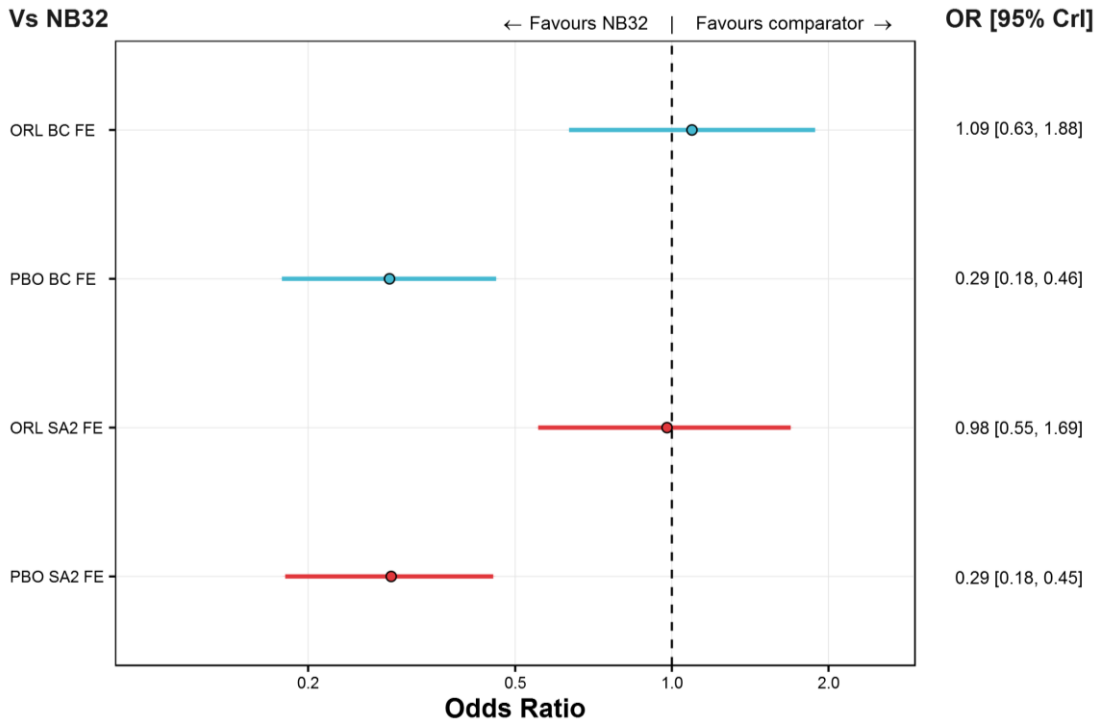
Model	Between-trial deviation (σ)	Treatment effects (d)	Trial baseline effects (μ)
Fixed effects	NA	$N(0,10000)$	$N(0,10000)$
Random effects	$U(0,5)$	$N(0,10000)$	$N(0,10000)$
Key: CFB, change from baseline; *N*, normal distribution; NA, not applicable; *U*, uniform distribution.			

Results

At least 5% reduction in weight

Results are presented as ORs with 95% CrI on the log scale. An OR less than one favours NB32 over orlistat or placebo. Figure 19 displays the results of the NMA and sensitivity analyses for ≥5% reduction in weight at 1 year for trials that specified patients with T2DM in the inclusion criteria. For the base case analysis, patients who receive orlistat have marginally higher odds of achieving ≥5% reduction in weight at 1 year than with NB32; however, this result is non-significant. Patients who receive placebo have significantly lower odds of achieving ≥5% reduction in weight compared to NB32 (OR: 0.29 [95% CrI: 0.18, 0.46]). The results of SA2, where trials that had lead-in periods were excluded, are similar to the base case analysis.

Figure 19: Forest plot for ≥5% reduction in weight (1 year) – T2DM – NMA results

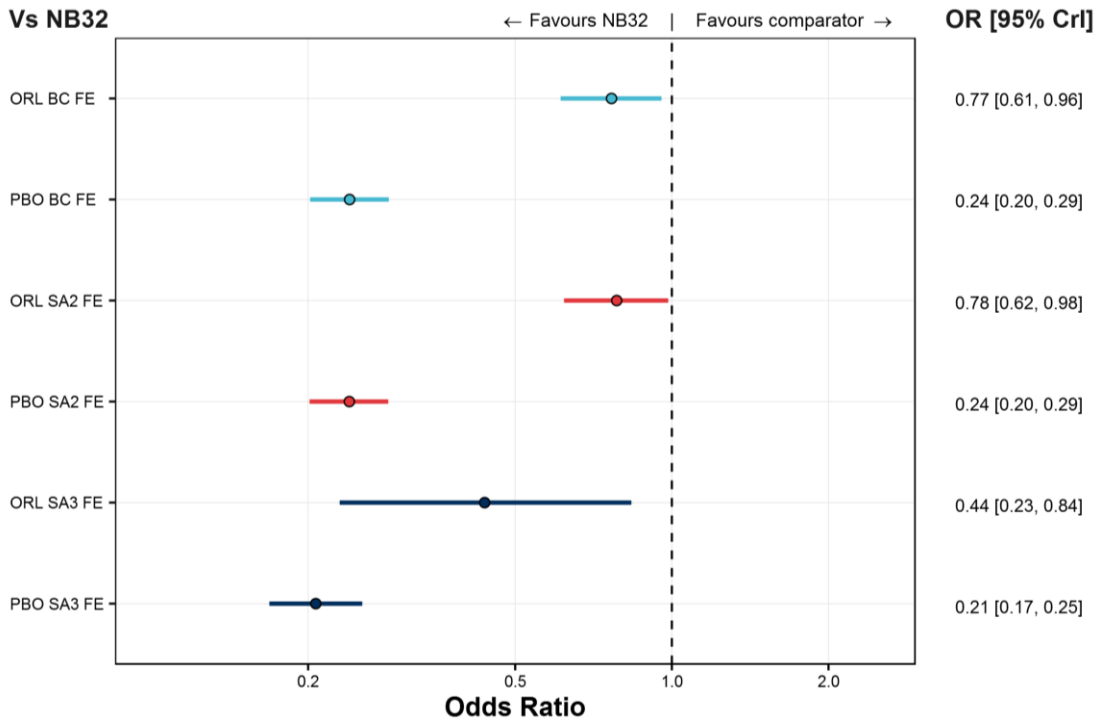


Key: BC, base case; CrI, credible interval; FE, fixed effects; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; OR, odds ratio; ORL, orlistat; PBO, placebo; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.
Notes: SA1 not performed due to insufficient data; SA3 not performed as repeat of the base case analysis; SA4 not performed as repeat of SA2.

Figure 20 displays the results of the NMA and sensitivity analyses for ≥5% reduction in weight at 1 year for trials that specified patients without T2DM in the inclusion

criteria. For the base case analysis, patients who receive orlistat have significantly lower odds of achieving ≥5% reduction in weight at 1 year than with NB32 (OR 0.77 [95% CrI: 0.61, 0.96]). Patients who receive placebo have significantly lower odds of achieving ≥5% reduction in weight compared to NB32 (OR: 0.24 [95% CrI: 0.20, 0.29]). The results of SA2 where trials that had lead-in periods were excluded are very similar to the base case analysis. In SA3, where trials with intensive behaviour modification were excluded, the odds of achieving ≥5% reduction while receiving orlistat was again significantly lower than NB32; however, the OR was much lower than the base case analysis (OR: 0.44 [95% CrI: 0.23, 0.84]).

Figure 20: Forest plot for ≥5% reduction in weight (1 year) – no T2DM – NMA results

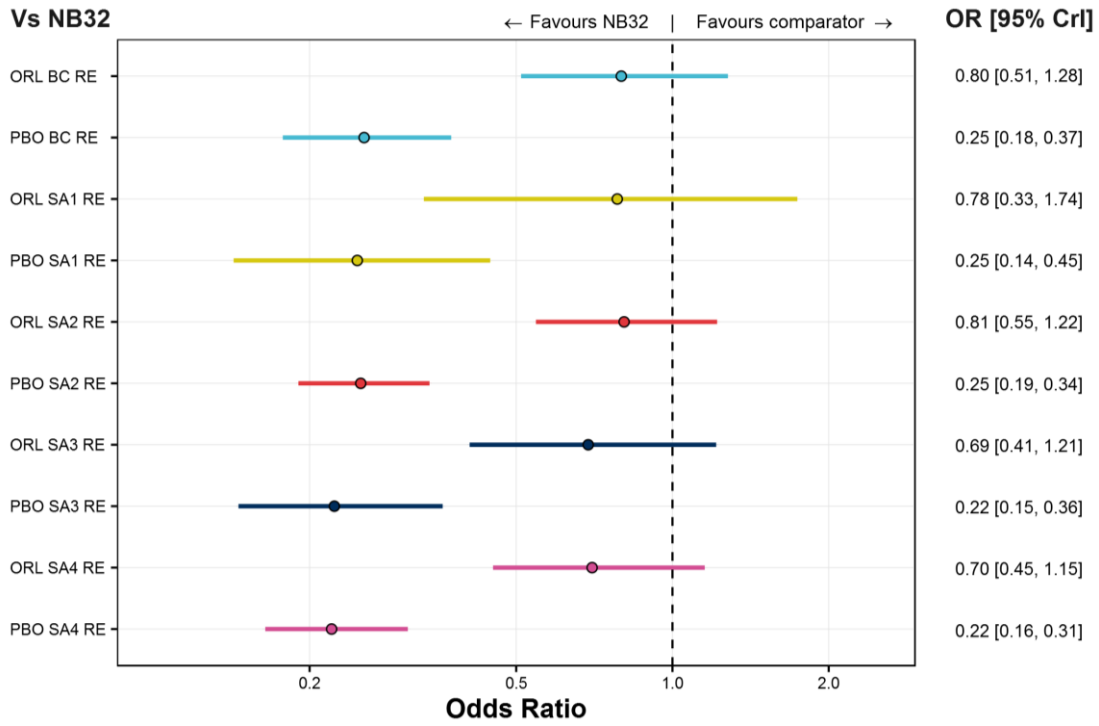


Key: BC, base case; CrI, credible interval; FE, fixed effects; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; OR, odds ratio; ORL, orlistat; PBO, placebo; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.
Notes: SA1 not performed as it is a repeat of the base case analysis; SA4 not performed due to insufficient data.

Figure **21** displays the results of the NMA and sensitivity analyses for ≥5% reduction in weight at 1 year for all trials. For the base case analysis, patients who receive orlistat have lower odds of achieving ≥5% reduction in weight at 1 year than with NB32; however, this result is non-significant. Patients who receive placebo have

significantly lower odds of achieving ≥5% reduction in weight compared to NB32 (OR: 0.25 [95% CrI: 0.18, 0.37]). The sensitivity analyses show that the results of SA1 and SA2 are similar to the base case analysis. The results of SA3 and SA4 produce similar results, which are slightly more favourable for NB32 compared to the base case analysis; however, the comparison of orlistat against NB32 is again non-significant.

Figure 21: Forest plot for ≥5% reduction in weight (1 year) – any T2DM – NMA results



Key: BC, base case; CrI, credible interval; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; OR, odds ratio; ORL, orlistat; PBO, placebo; RE, random effects; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.

The results of the ≥5% reduction in weight at 1 year outcome indicate patients without T2DM have significantly lower odds of achieving ≥5% reduction in weight while receiving orlistat or placebo compared to NB32. While in patients who have T2DM the analysis indicates that patients receiving NB32 have comparable odds of achieving ≥5% reduction in weight compared to orlistat, and placebo has significantly lower odds compared to NB32.

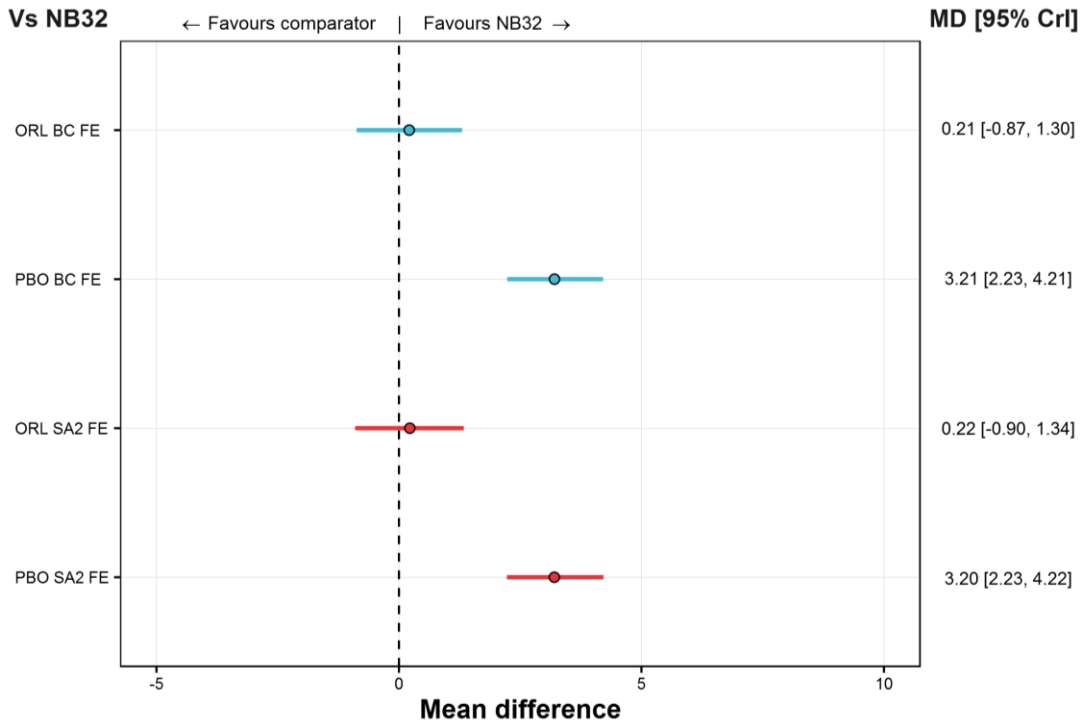
The sensitivity analysis results indicate the following:

- The removal of trials that have a high proportion of patients who have comorbidities (SA1) produce similar results to the base case analysis.
- The removal of trials with lead-in periods (SA2) produces similar results to the base case analysis.
- The exclusion of 'intensive' behaviour modification in SA3 and SA4 produces results which are slightly more favourable for NB32 than the base case analysis; the additional exclusion of studies with lead-in periods in SA4 produces similar results compared to SA3.

Percentage weight change from baseline

Results are presented as MDs with 95% CrIs on a linear scale. A MD of >0 favours NB32 over orlistat or placebo and indicates greater % weight reduction. Figure 22 displays the results of the NMA and sensitivity analyses for mean % weight change from baseline at 1 year for trials that specified patients with T2DM in the inclusion criteria. For the base case analysis, patients who receive orlistat have marginally lower % weight reduction at 1 year compared to NB32; however, this result is non-significant. Patients who receive placebo have significantly lower % reduction in weight than NB32 patients (MD: 3.21 [95% CrI, 2.23, 4.21]). The results of SA2 where trials with lead-in periods have been removed produced similar results to the base case analysis.

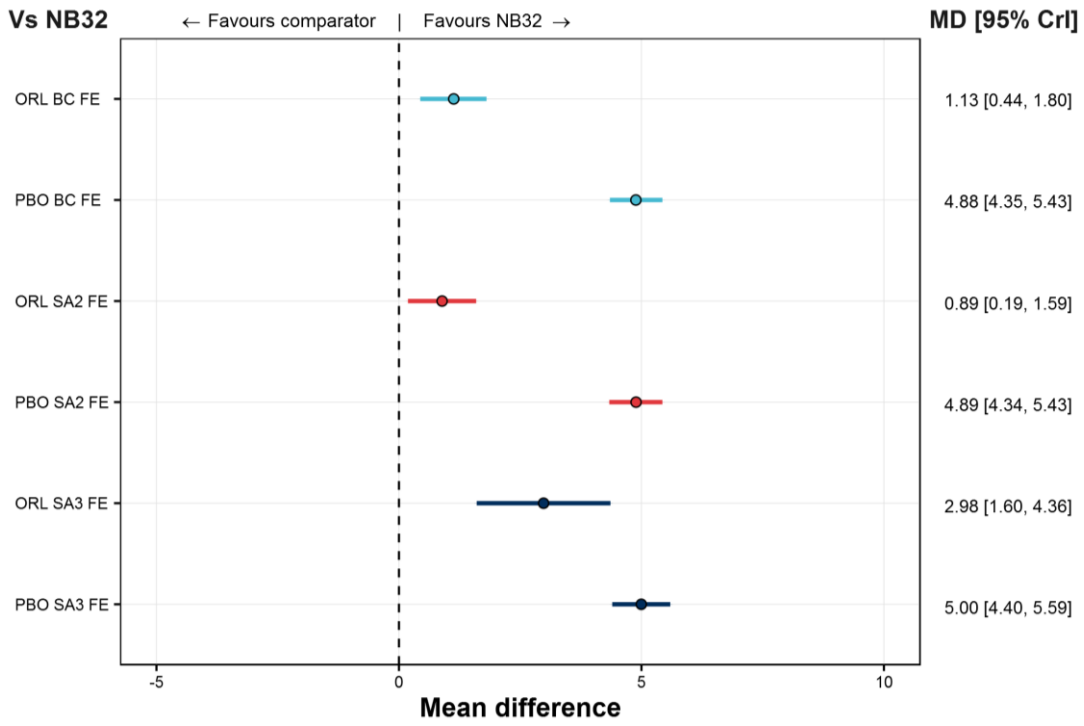
Figure 22: Forest plot for mean % weight CFB (1 year) – T2DM – NMA results



Key: BC, base case; CFB, change from baseline; CrI, credible interval; FE, fixed effects; MD, mean difference; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; ORL, orlistat; PBO, placebo; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.
Notes: SA1 not performed as insufficient data; SA3 not performed as repeat of base case analysis; SA4 not performed as repeat of SA2.

Figure 23 displays the results of the NMA and sensitivity for mean % weight change from baseline at 1 year for trials which specified patients without T2DM in the inclusion criteria. For the base case analysis, patients who receive orlistat have a significantly lower % weight reduction at 1 year compared to NB32 (MD: 1.13 [95% CrI: 0.44, 1.80]). The results of SA2 suggest that the MD of % weight change from baseline is marginally reduced when trials with lead-in periods are excluded. In SA3, the mean % weight reduction was again significantly lower than NB32; however, the MD was greater than the base case analysis (MD: 2.98 [95% CrI: 1.60, 4.36]).

Figure 23: Forest plot for mean % weight CFB (1 year) – no T2DM – NMA results

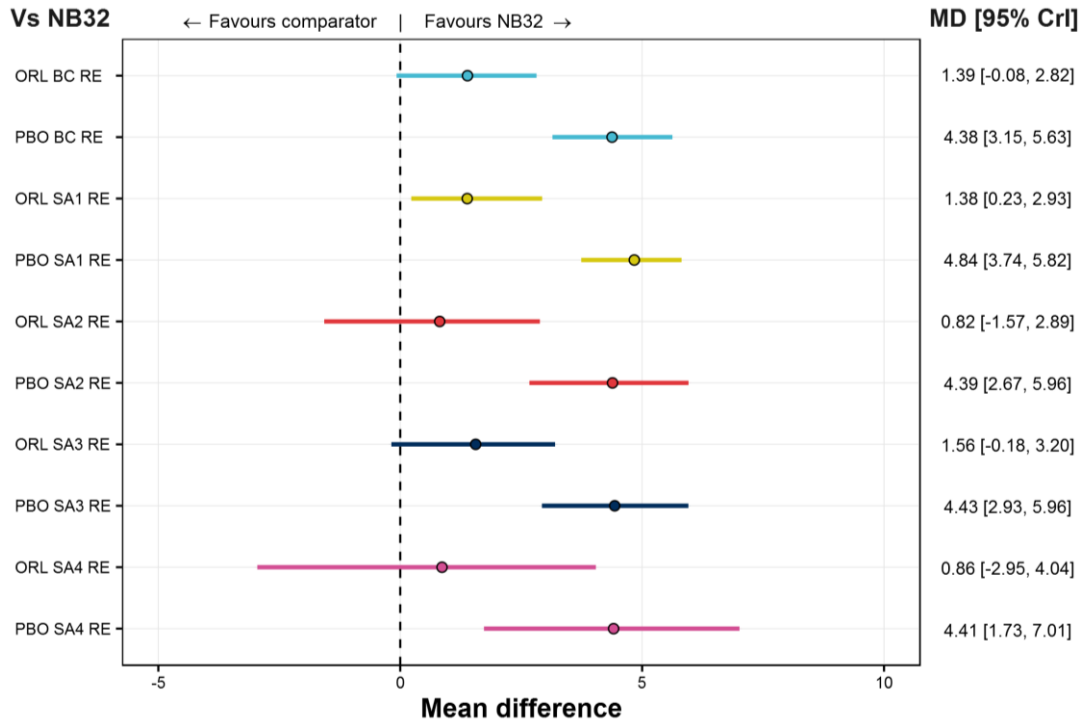


Key: CFB, change from baseline; CrI, credible interval; FE, fixed effects; MD, mean difference; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; ORL, orlistat; PBO, placebo; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.
Notes: SA1 not performed as repeat of base case analysis; SA4 not performed as insufficient data available.

Figure 24 displays the results of the NMA and sensitivity analyses for mean % weight change from baseline at 1 year for all trials. For the base case analysis, patients who receive orlistat have a lower % weight reduction at 1 year compared to NB32; however, this result is non-significant. The results of SA1, where trials with a high proportion of patients with comorbidities are excluded, produce a similar result for the comparison of orlistat versus NB32 compared to the base case analysis; however, less uncertainty is observed, which produces a significant result (MD: 1.38 [95% CrI: 0.23, 2.93]). SA4 produces results that are similar to the base case analysis, suggesting that intensive behaviour modification therapy has little effect on the relative % weight change from baseline. This is seen further when comparing SA2 and SA4, which produce similar results despite the additional exclusion of intensive behaviour modification therapy trials in SA4. The results of SA2 and SA4

also produce a slightly lower MD than the base case analysis; however, the comparison of orlistat versus NB32 still favours NB32.

Figure 24: Forest plot for mean % weight CFB (1 year) – any T2DM – NMA results



Key: BC, base case; CFB, change from baseline; CrI, credible interval; MD, mean difference; NB32, naltrexone 32mg plus bupropion; NMA, network meta-analysis; ORL, orlistat; PBO, placebo; RE, random effects; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.

The results of the mean % weight change from baseline at 1 year outcome indicate that patients without T2DM experience significantly greater % weight loss while receiving NB32 compared to orlistat and placebo. Whist in patients who have T2DM the analysis indicates that patients treated with NB32 may have similar % weight loss compared with orlistat, and significantly greater % weight loss compared to placebo.

The sensitivity analysis results indicate the following:

- The removal of trials that have a high proportion of patients who have comorbidities (SA1) produces similar results to the base case; however, less uncertainty is seen in the estimate.

- The removal of trials with lead-in periods (SA2 compared to base case analysis and SA4 compared to SA3) is seen to generally reduce the MD of the % weight change from baseline for the comparison of orlistat versus NB32; however, the results still favour NB32.
- The removal of trials with intensive behaviour modification (SA3) is seen to increase the MD of the 5% weight change from baseline; however, this increase is only marginal in the any-T2DM analysis. The MD compared to placebo in SA3 is consistent with the estimate from the base case analysis which supports the assumption that the effects of the additional behaviour modification is additive.

Goodness-of-fit

Table 38 and Table 39 present the mean total residual deviance for each model by outcome to assess the goodness-of-fit. Most models show a reasonably good fit to the data; however, the two T2DM models (base case and SA2) for % weight change from baseline displayed a significantly poor fit. In the forest plots for these two analyses (Figure 37 to Figure 40 in Appendix 13), the Derosa 2010 results look different to the other orlistat trials.[74] As fixed effects estimates have been used in these two analyses, the heterogeneity between the orlistat trials may not be adequately captured when fitting the model, which is likely to contribute to the poor model fit. Analysis excluding the Derosa 2010 trial was not considered, as the removal of this study is likely to bias results in favour of NB32; the Derosa 2010 trial produced the largest point estimate in favour of orlistat.

Table 38: Goodness-of-fit for each model – ≥5% reduction in weight

Analysis	T2DM population	Number of unique data points	Total residual deviance (mean)	p-value[a]
Base case (FE)	T2DM	8	9.6	0.295
SA1 (FE)		NA		
SA2 (FE)		6	5.1	0.537
SA3 (FE)		NA		
SA4 (FE)		NA		
Base case (FE)	Non-T2DM	10	17.4	0.066
SA1 (FE)		NA		

Analysis	T2DM population	Number of unique data points	Total residual deviance (mean)	p-value[a]
SA2 (FE)		8	14.4	0.072
SA3 (FE)		6	5.1	0.535
SA4 (FE)		NA		
Base case (RE)	Any T2DM	24	26.1	0.349
SA1 (RE)		12	12.8	0.384
SA2 (RE)		16	16.1	0.445
SA3 (RE)		22	22.0	0.458
SA4 (RE)		12	10.5	0.568

Key: FE, fixed effects NA, not applicable; RE, random effects; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.
Notes: [a], Test for differences between number of data points and total residual deviance is based on a Chi squared test.

Table 39: Goodness-of-fit for each model – % weight CFB

Analysis	T2DM population	Number of unique data points	Total residual deviance (mean)	p-value[a]
Base case (FE)	T2DM	10	23.4	0.009[b]
SA1 (FE)		NA		
SA2 (FE)		8	22.4	0.004[b]
SA3 (FE)		NA		
SA4 (FE)		NA		
Base case (FE)	Non-T2DM	16	20.8	0.186
SA1 (FE)		NA		
SA2 (FE)		8	7.5	0.484
SA3 (FE)		12	9.4	0.668
SA4 (FE)		NA		
Base case (RE)	Any T2DM	40	40.5	0.447
SA1 (RE)		22	20.9	0.529
SA2 (RE)		18	19.7	0.349
SA3 (RE)		36	36.8	0.433
SA4 (RE)		14	15.2	0.365

Key: CFB, change from baseline; FE, fixed effects; NA, not applicable; RE, random effects; SA, sensitivity analysis; T2DM, Type 2 diabetes mellitus.
Notes: [a], Test for differences between number of data points and total residual deviance is based on a Chi squared test

Between-trial heterogeneity

The statistical measure of the between-trial heterogeneity for the any-T2DM models for each outcome are presented in Table 40 and Table 41. Between-trial heterogeneity is not presented for the T2DM and no-T2DM models as random effects were not fitted. For the ≥5% reduction in weight, the between-trial deviation indicates mild heterogeneity for all analyses. For both outcomes, the 95% credible intervals are reasonably wide, suggesting some uncertainty around the true amount of heterogeneity.

Table 40: Between-trial heterogeneity for each model – ≥5% reduction in weight

Analysis	Between-trial deviation, median (95% CrI)
Base case	0.27 (0.07, 0.61)
SA1	0.31 (0.05, 1.12)
SA2	0.17 (0.01, 0.52)
SA3	0.27 (0.03, 0.69)
SA4	0.11 (0.00, 0.60)
Key: CrI, credible interval; SA, sensitivity analysis.	

Table 41: Between-trial heterogeneity for each model – % weight CFB

Analysis	Between-trial deviation, median (95% CrI)
Base case	1.07 (0.64, 1.84)
SA1	0.53 (0.03, 1.72)
SA2	1.30 (0.56, 3.02)
SA3	1.12 (0.61, 2.04)
SA4	1.77 (0.65, 4.18)
Key: CFB, change from baseline; CrI, credible interval; SA, sensitivity analysis.	

Consistency of direct and indirect evidence

Consistency could not be checked as there were no closed loops within the network.

4.11 Non-randomised and non-controlled evidence

Non-RCT evidence was not formally considered as part of comparative efficacy or cost-effectiveness assessments as RCT data were available for the intervention and comparators of interest to the decision problem.

4.12 Adverse reactions

NB32 was generally well tolerated, with readily manageable AEs consistent with the well-established safety profiles of naltrexone and bupropion. The pattern of treatment-emergent adverse events (TEAEs) of obese patients with T2DM was similar to that in non-diabetic patients.

TEAEs were defined as events that first occurred or worsened during double-blind treatment (i.e. a new event or an exacerbation of a pre-existing condition) with an onset date after study drug administration and within 7 days of the last confirmed dose date. AEs with an onset date before the first dose of study drug were recorded under medical history.

Safety data are presented for the safety analysis set, defined as all randomised patients who were administered at least one tablet of study treatment and had at least one investigator contact/assessment at any time after the start of study treatment, regardless of whether they discontinued the study. Patients were grouped in the safety analysis set according to which study treatment was administered on the first day of treatment following randomisation.

Of note, safety outcomes for blood pressure and pulse rate were generally comparable to those presented in Section 4.7.

COR-I study: Safety profile

Overall, more patients in the NB16 and NB32 treatment groups experienced at least one TEAE and discontinued the study drug due to an AE compared to patients treated with placebo. In addition, drug-related TEAEs were higher in NB16 and NB32 patients (57.1% and 58.6% respectively) compared to placebo (29.3%). Most TEAEs were considered mild or moderate in severity, with incidences of severe TEAEs <10% in all treatment groups (NB32: 8.9%; NB16: 9.7%; placebo: 6.0%).

There were similarly low incidences of treatment-emergent serious adverse events (TESAEs) in all treatment groups (1.6% for NB16 and NB32 treatment groups and 1.4% for placebo). One patient in the NB32 treatment group died during the study. This was due to a myocardial infarction and was considered unlikely to be related to study drug.

Table 42: Summary of safety data from COR-I, safety analysis set

	NB16 (n=569)	NB32 (n=573)	Placebo (n=569)
All TEAEs, n (%)	455 (80.0)	476 (83.1)	390 (68.5)
Drug-related TEAEs, n (%)	325 (57.1)	336 (58.6)	167 (29.3)
Severe TEAEs, n (%)	55 (9.7)	51 (8.9)	34 (6.0)
TESAEs, n (%)	9 (1.6)	9 (1.6)	8 (1.4)
DC due to AEs, n (%)	122 (21.4)	112 (19.5)	56 (9.8)
Deaths, n (%)	0	1 (0.2)	0
Key: AE, adverse event; DC, discontinuation; TEAE, treatment emergent adverse event; TESAE, treatment emergent serious adverse event. **Source:** Greenway et al. 2010[16]; Orexigen, 2010[56]			

AEs in the NB groups were most frequently GI in nature. The most common of these, nausea, was generally mild to moderate in intensity, transient, and did not result in discontinuation for most participants who reported it (Table 43). Nausea was typically first reported during dose escalation in the experimental groups; the rate of onset seemed to plateau shortly after reaching full dose and then was similar to the rate reported in the placebo group.

A total of 171 patients (29.8%) in the NB32 group had nausea; however, only 36 patients (6.3%) discontinued because of this. Other AEs leading to discontinuation included headache (0.9%) and depression (0.2%).

Table 43: Select AE data from COR-I, safety analysis set

	NB16 (n=569)	NB32 (n=573)	Placebo (n=569)
Any AE, n (%)			
Nausea	155 (27.2)	171 (29.8)	30 (5.3)
Headache	91 (16.0)	79 (13.8)	53 (9.3)
Constipation	90 (15.8)	90 (15.7)	32 (5.6)
Upper respiratory tract infection	49 (8.6)	57 (9.9)	64 (11.2)

	NB16 (n=569)	NB32 (n=573)	Placebo (n=569)
Dizziness	44 (7.7)	54 (9.4)	15 (2.6)
Insomnia	36 (6.3)	43 (7.5)	29 (5.1)
Vomiting	36 (6.3)	56 (9.8)	14 (2.5)
Sinusitis	34 (6.0)	30 (5.2)	34 (6.0)
Dry mouth	42 (7.4)	43 (7.5)	11 (1.9)
Nasopharyngitis	32 (5.6)	29 (5.1)	31 (5.4)
Diarrhoea	31 (5.4)	26 (4.5)	28 (4.9)
Hot flush	13 (2.3)	30 (5.2)	7 (1.2)
Psychiatric AE, n (%)			
Insomnia	36 (6.3)	43 (7.5)	29 (5.1)
Anxiety	12 (2.1)	9 (1.6)	12 (2.1)
Depression	9 (1.6)	3 (0.5)	6 (1.1)
Any AE leading to DC, n (%)	122 (21.4)	112 (19.5)	56 (9.8)
Gastrointestinal disorders	42 (7.4)	48 (8.4)	9 (1.6)
Nausea	26 (4.6)	36 (6.3)	2 (0.4)
Nervous system disorders	30 (5.3)	19 (3.3)	15 (2.6)
Dizziness	13 (2.3)	7 (1.2)	3 (0.5)
Headache	9 (1.6)	5 (0.9)	4 (0.7)
Psychiatric disorders	13 (2.3)	12 (2.1)	11 (1.9)
Depression	6 (1.1)	1 (0.2)	2 (0.4)

Key: AE, adverse event; DC, discontinuation.
Source: Greenway et al. 2010[16]

Given the relatively common overlap between depression and obesity, and due to the withdrawal of previously approved weight management agents due to serious psychiatric side effects, assessment of such effects is an important safety consideration which was measured within the four pivotal studies.[91] The IDS-SR was used as a screening tool to exclude individuals with depression from enrolling in the studies and also used to assess changes in mood or depressive symptoms over the course of the studies. This self-rated instrument was used as an assessment of both efficacy (presented in Section 4.7) and safety. For the safety analysis, treatment-emergent depressive or anxiety symptoms on the IDS-SR were defined as a score of

≥2 on items 5 (sadness), 6 (irritability), 7 (anxiety/tension) or 18 (suicidality), or a total score ≥25 (or ≥30 for patients with a total score ≥25 at screening).

Changes in depression related symptoms were monitored at each study visit. For all randomised patients, total IDS-SR score (range of possible scores: 0 to 84) was generally low at baseline with a median of 5 (range: 0–36) for the NB32 treatment group.

Results of the IDS-SR total score and individual depressive and anxiety symptom items during double-blind treatment were similar across the treatment groups.[56] A total of three patients in the NB32 group and one patient in the placebo group had treatment-emergent post-baseline scores of ≥2 on item #18 (suicidality) of the IDS-SR questionnaire. Three additional patients, one in the NB16 treatment group and two receiving placebo, had a score of ≥2 more than 24 hours post-last dose. Of these seven patients, five completed the study (two in each of the placebo and NB32 groups, and one in the NB16 group).[56]

Throughout the trial, a total of three CV serious adverse events (SAEs) were reported: one pericardial effusion (placebo group), one cardiac failure (NB32 group) and one death due to acute myocardial infarction (a patient with multiple CV risk factors assigned to treatment with NB32). Investigators did not regard these as related to study drug.

Adverse events during the drug discontinuation phase for NB16 (7.7% for sudden and 9.2% for tapered) and NB32 (8.9% sudden and 9.5% tapered) groups occurred at similar rates to placebo (8.5%), regardless of discontinuation method. No patients experienced an SAE during the discontinuation phase. One patient each in the NB16 (tapered), NB32 (tapered), and placebo groups experienced a discontinuation-emergent AE (DEAE) of severe intensity compared to no patients in the sudden discontinuation groups for NB16 and NB32. The majority of DEAEs were considered not related to study drug, regardless of the treatment group or discontinuation method.

COR-II study: Safety profile

Overall, NB was associated with a greater incidence of TEAEs than placebo and more patients in the NB groups discontinued treatment because of an TEAE (Table 44) particularly early in the trial.[17] Most TEAEs were considered mild or moderate in

severity by investigators, while there was a greater incidence of severe TEAEs in the NB group (11.1%) compared to the placebo group (6.7%). In addition, more patients in the NB group experienced a TEAE considered drug-related by investigators (63.5% vs 38.4% in placebo-treated patients).

Although TESAEs were greater in the NB32 group than placebo-treated patients, rates were low in both groups (2.1% vs 1.4%, respectively[17]).

Table 44: Summary of safety data from COR-II, safety analysis set

	NB32/48 (n=992)	**Placebo (n=492)**
All TEAEs, n (%)	852 (85.9)	370 (75.2)
Drug-related TEAEs, n (%)	630 (63.5)	189 (38.4)
Severe TEAEs, n (%)	110 (11.1)	33 (6.7)
TESAEs, n (%)	21 (2.1)	7 (1.4)
DC due to AEs, n (%)	241 (24.3)	68 (13.8)
Deaths, n (%)	0	0
Key: AE, adverse event; DC, discontinuation; SAE, serious adverse event; TEAE, treatment emergent adverse event; TESAE, treatment emergent serious adverse event. **Source:** Apovian et al. 2013[17]; Orexigen, 2010[62]		

The most frequent TEAEs were nausea, headache and constipation (Table 45). These events were mostly mild to moderate and did not result in discontinuation in most patients who experienced them. Most nausea events occurred during the dose escalation period and were transient. There was one event of passive suicidal ideation in an NB32-treated patient; symptoms resolved following study drug discontinuation.

As shown in Table 45, the most common reason for discontinuation was nausea, which was reported in 6.0% of patients in the NB treatment group compared to 0.2% of placebo-treated patients ($p<0.05$). More patients in the NB treatment group also discontinued due to headache ($p<0.05$ vs placebo) and depression (0.5% of NB treatment patients compared to 1.2% of placebo treated patients).

Table 45: Select AE data from the COR-II study, safety analysis set

	NB32/48 (n=992)	**Placebo (n=492)**
Any AE, n (%)	852 (85.9)	370 (75.2)
Nausea	290 (29.2)	34 (6.9)

	NB32/48 (n=992)	Placebo (n=492)
Constipation	189 (19.1)	35 (7.1)
Headache	174 (17.5)	43 (8.7)
Insomnia	97 (9.8)	33 (6.7)
Dry mouth	90 (9.1)	13 (2.6)
Upper respiratory tract infection	86 (8.7)	55 (11.2)
Vomiting	84 (8.5)	10 (2.0)
Nasopharyngitis	82 (8.3)	40 (8.1)
Dizziness	68 (6.9)	18 (3.7)
Diarrhoea	55 (5.5)	18 (3.7)
Sinusitis	51 (5.1)	35 (7.1)
Arthralgia	38 (3.8)	28 (5.7)
Bronchitis	14 (1.4)	25 (5.1)
Any psychiatric AE, n (%)	205 (20.7)	75 (15.2)
Insomnia	97 (9.8)	33 (6.7)
Anxiety	48 (4.8)	21 (4.3)
Depression	13 (1.3)	8 (1.6)
Sleep disorder	11 (1.1)	4 (0.8)
Any AE leading to DC, n (%)	241 (24.3)	68 (13.8)
Nausea	60 (6.0)	1 (0.2)
Headache	26 (2.6)	4 (0.8)
Depression	5 (0.5)	6 (1.2)
Key: AE, adverse event; DC, discontinuation; NB32, naltrexone 32mg plus bupropion **Source:** Apovian et al. 2013[17]		

As discussed for the COR-I study, the IDS-SR was used as a screening tool and as an assessment of safety. A total of two patients in the NB32 group and one patient in the placebo group had a score of ≥2 on item #18 (suicidality) of the IDS-SR questionnaire during double-blind treatment. None of these patients were receiving relevant concomitant medications at baseline. Of these, one NB32-treated patient had TEAEs of depression and suicidal ideation that resulted in study drug discontinuation[17], while the other two patients completed the study.[56] NB was not associated with increased incidence of treatment-emergent symptoms of depression or other mood-related AEs.

There was one myocardial infarction in an NB-treated patient with active coronary artery disease, angina pectoris, hyperlipidaemia, and hypertension. One seizure was reported for an NB-treated patient with no history of seizures. There were no clinically significant effects of NB on laboratory measures or electrocardiography.

COR-BMOD study: Safety profile

Overall, a greater proportion of patients in the NB32 + BMOD treatment group experienced at least one TEAE compared to patients in the placebo group (93.7% compared to 88.0%, respectively [**Error! Not a valid bookmark self-reference.**]). Similarly, more patients in the NB32 + BMOD group discontinued the study due to an AE (25.7% vs 12.5% in the placebo group).[63]

Most TEAEs were considered mild or moderate in severity although 16.8% of patients in the NB32 group had a severe TEAE compared to 7.5% of patients receiving placebo.[63] In addition, more patients in the NB32 group experienced at least one serious TEAE, but rates were low in both groups (3.8% compared to 0.5% in the placebo group).[63]

Table 46: Summary of safety data from the COR-BMOD study, safety analysis set

	NB32 + BMOD (n=584)	Placebo + BMOD (n=200)
All TEAEs, n (%)	547 (93.7)	176 (88.0)
Drug-related TEAEs, n (%)	447 (76.5)	108 (54.0)
Severe TEAEs, n (%)	98 (16.8)	15 (7.5)
TESAEs, n (%)	22 (3.8)	1 (0.5)
DC due to AEs, n (%)	150 (25.7)	25 (12.5)
Deaths, n (%)	0	0

Key: AE, adverse event; DC, discontinuation; SAE, serious adverse event; TEAE, treatment emergent adverse event; TESAE, treatment emergent serious adverse event.
Source: Orexigen, 2010[62]

Table 47 presents AEs that occurred in ≥5% of patients in either treatment group. Nausea was the most frequent AE, with 34.1% of participants treated with NB32 + BMOD reporting at least one event, compared to 10.5% for placebo + BMOD ($p<0.001$). Nausea was mostly mild to moderate in intensity and occurred primarily during the first 4 weeks of the study (coinciding with drug titration), with a median

duration of 10 days with NB32 + BMOD and 12 days with placebo + BMOD. Constipation, dizziness, dry mouth, tremor, upper abdominal pain, and tinnitus also occurred more often in the NB32 + BMOD group than in placebo + BMOD.

Two SAEs occurred in the NB32 + BMOD group that were considered possibly related to study drug. Both involved cholecystitis in patients who had experienced marked weight loss (>15kg). Both patients resumed blinded therapy after successful surgical treatment.

As shown in Table 47, nausea was the most frequent AE that resulted in study drug discontinuation (4.6% in the NB32 + BMOD group vs 0% in the placebo + BMOD group; p<0.001). Other frequent AEs that resulted in study drug discontinuation in >0.5% of NB32 + BMOD treated patients included urticarial, anxiety, disturbance in attention, headache, increase in blood pressure, dizziness and vomiting. However, in none of these cases did the incidence of discontinuation for a specific AE have a p value <0.05 for NB32 + BMOD versus placebo + BMOD.

In nearly 10% of NB32 + BMOD treated patients who discontinued due to an AE, the AEs contributing to discontinuation were of a wide variety that occurred at a very low frequency (i.e. ≤0.3%). Among the 15.2% of patients in the NB32 + BMOD group who discontinued the study drug in the first month due to an AE, nausea was the most common event, accounting for 2.9% of discontinuations, compared with 0% in the placebo + BMOD group (p=0.010). In general, the remaining study drug discontinuations due to an AE in the first month were attributable to the same AEs shown in Table 47.

Table 47: Key AE data from the COR-BMOD study, safety analysis set

	NB32 + BMOD (n=584)	**Placebo + BMOD (n=200)**	**p-value**
Any AEs, n (%)			
Nausea	199 (34.1)	21 (10.5)	<0.001
Headache	139 (23.8)	35 (17.5)	0.076
Constipation	141 (24.1)	28 (14.0)	0.003
Dizziness	85 (14.6)	9 (4.5)	<0.001
Vomiting	64 (11.0)	13 (6.5)	0.074
Insomnia	51 (8.7)	12 (6.0)	0.291
Dry mouth	47 (8.0)	6 (3.0)	0.014

	NB32 + BMOD (n=584)	Placebo + BMOD (n=200)	p-value
Anxiety	30 (5.1)	7 (3.5)	0.441
Tremor	34 (5.8)	2 (1.0)	0.003
Upper abdominal pain	32 (5.5)	3 (1.5)	0.017
Tinnitus	31 (5.3)	1 (0.5)	0.001
Any psychiatric AEs, n (%)			
Insomnia	51 (8.7)	12 (6.0)	0.291
Anxiety	30 (5.1)	7 (3.5)	0.441
Sleep disorder	14 (2.4)	6 (3.0)	0.610
Depressed mood	11 (1.9)	8 (4.0)	0.110
Abnormal dreams	8 (1.4)	4 (2.0)	0.514
Middle insomnia	6 (1.0)	2 (1.0)	1.000
Tension	7 (1.2)	1 (0.5)	0.687
Depression	2 (0.3)	5 (2.5)	0.014
Stress	3 (0.5)	4 (2.0)	0.074
Dissociation	6 (1.0)	0 (0)	0.347
Any AEs resulting in DC, n (%)			
Nausea	27 (4.6)	0 (0)	<0.001
Urticaria	10 (1.7)	1 (0.5)	0.306
Anxiety	7 (1.2)	3 (1.5)	0.721
Disturbance in attention	6 (1.0)	0 (0)	0.347
Headache	5 (0.9)	1 (0.5)	1.000
Blood pressure increased	4 (0.7)	0 (0)	0.577
Dizziness	4 (0.7)	0 (0)	0.577
Vomiting	4 (0.7)	0 (0)	0.577
Depressed mood	3 (0.5)	1 (0.5)	1.000
Feeling abnormal	3 (0.5)	1 (0.5)	1.000
Abdominal pain	3 (0.5)	0 (0)	0.574
Upper abdominal pain	3 (0.5)	0 (0)	0.574
Disorientation	3 (0.5)	0 (0)	0.574
Dissociation	3 (0.5)	0 (0)	0.574
Feeling jittery	3 (0.5)	0 (0)	0.574
Insomnia	3 (0.5)	0 (0)	0.574
Rash	3 (0.5)	0 (0)	0.574

	NB32 + BMOD (n=584)	Placebo + BMOD (n=200)	p-value
Key: AE, adverse event; BMOD, intensive behaviour modification; DC, discontinuation **Source:** Wadden et al. 2010[18]			

With one exception, there were no differences between groups in the 10 most frequently observed psychiatric AEs.[18] However, depression occurred more frequently in the placebo + BMOD group (2.5% vs 0.3% in the NB32 group; p<0.014).[18] A total of three patients (two in the NB32 group and one in the placebo group) had a score of ≥2 in item #18 (suicidality) of the IDS-SR questionnaire.[63] All three patients completed the study.[63]

COR-DM study: safety profile

Overall, more patients in the NB32 treatment group experienced at least one TEAE during double-blind treatment (90.4% vs 85.2% in the placebo group). Discontinuation due to an AE was relatively high (29.4% vs 15.4% [Table 48]). A similar percentage of patients in the NB32 group and placebo group experienced at least one TESAE (3.9% vs 4.7%).[19] Most TEAEs were considered mild or moderate in severity by investigators. A greater incidence of patients in the NB32 group compared to placebo-treated patients experienced a severe TEAE (18.3% vs 11.2% respectively) and drug-related TEAEs (71.5% vs 33.7%). No patients died during the study.[64]

Table 48: Summary of safety data from the COR-DM study, safety analysis set

	NB32 (n=333)	Placebo (n=169)
All TEAEs, n (%)	301 (90.4)	144 (85.2)
Drug-related TEAEs, n (%)	238 (71.5)	57 (33.7)
Severe TEAEs, n (%)	61 (18.3)	19 (11.2)
TESAEs, n (%)	13 (3.9)	8 (4.7)
DC due to AEs, n (%)	98 (29.4)	26 (15.4)
Deaths, n (%)	0	0
Key: AE, adverse event; DC, discontinuation; SAE, serious adverse event; TEAE, treatment emergent adverse event; TESAE, treatment emergent serious adverse event. **Source:** Hollander et al. 2013[19]; Orexigen, 2009[64]		

The most common AEs that were more prevalent in the NB-treated patients were nausea, constipation, vomiting and diarrhoea (Table 49). Nausea led to withdrawal in 9.6% of NB32-treated patients, with the vast majority (28 out of 32) of these withdrawals occurring as the result of nausea with an onset during the first 4 weeks of treatment.[19]

AEs that led to medication discontinuation during the first 4 weeks of treatment were the primary reason that relatively fewer randomised NB32 patients were included in the mITT population. Nausea occurred more frequent in NB32-treated patients taking metformin at baseline (46.2%) compared with those not on metformin (28.2%).[19] The incidence of patients with SAEs was low (3.9% for NB and 4.7% for placebo) and similar to that previously reported for patients without T2DM.[16-18]

The most common reason for discontinuation was nausea (9.6% of NB32 patients compared to 0% of placebo treated patients). Other common reasons for discontinuations in the NB32 group included vomiting (3%), headache (1.8%) and depression (0.6%). A similar proportion of patients in both group discontinued due to diabetes-related complications.

Table 49: Key AE data from the COR-DM study, safety analysis set

	NB32 (n=333)	**Placebo (n=169)**	**p-value**
Any AE, n (%)			
Nausea	141 (42.3)	12 (7.1)	<0.001
Constipation	59 (17.7)	12 (7.1)	0.001
Vomiting	61 (18.3)	6 (3.6)	<0.001
Diarrhoea	52 (15.6)	16 (9.5)	0.072
Headache	46 (13.8)	15 (8.9)	0.115
Dizziness	390 (11.7)	9 (5.3)	0.024
Insomnia	37 (11.1)	9 (5.3)	0.034
Nasopharyngitis	28 (8.4)	23 (13.6)	0.085
Hypertension	33 (9.9)	7 (4.1)	0.024
Upper respiratory tract infection	26 (7.8)	16 (9.5)	0.609
Hypoglycaemia	25 (7.5)	12 (7.1)	1.000
Tremor	22 (6.6)	4 (2.4)	0.054
Dry mouth	21 (6.3)	5 (3.0)	0.137
Anxiety	18 (5.4)	2 (1.2)	0.027

Upper abdominal pain	17 (5.1)	3 (1.8)	0.091
Discontinuations due to AEs, n (%)			
Nausea	32 (9.6)	0 (0)	<0.001
Vomiting	10 (3)	0 (0)	0.019
Headache	6 (1.8)	0 (0)	0.102
Depression	2 (0.6)	3 (1.8)	0.341
Diabetes	1 (0.3)	2 (1.2)	0.263
Hyperglycaemia	0 (0)	2 (1.2)	0.113
Key: AE, adverse event; NB32, naltrexone 32g plus bupropion **Source:** Hollander et al. 20213[19]; Orexigen, 2009[64]			

Results for the IDS-SR total score and individual depressive items during double-blind treatment were similar across the treatment groups. Anxiety symptoms were higher with NB32 treatment than placebo.[64] One placebo-treated patient had a TEAE of suicidal ideation but did not have a suicidality score ≥2 at any time. No other subject had a TEAE related to suicidality.[64]

The NB-CVOT Study

A summary of safety data from the NB-CVOT study is presented in Table 50. Only SAEs and AEs leading to study drug discontinuation were collected. More patients in the NB32 group experienced events that were considered by the investigator to be study drug–related (22.0% vs 3.9% with placebo). In both groups, most TEAEs leading to discontinuation were considered mild or moderate in intensity. TESAEs, (defined as any AE occurring at any dose of study drug that resulted in death, life-threatening adverse drug experience, inpatient hospitalisation or prolongation of existing hospitalisation, persistent of significant disability or incapacity, important medical events or congenital anomaly or birth defect) were reported for 849 patients (9.5%) overall, 10.4% in the NB32 group and 8.7% in the placebo group. The greatest treatment-group difference was seen in the GI disorders, which were reported more often with NB32. The percentage of patients with study drug–related SAEs was 0.3% and 0.2% for NB32 and placebo, respectively. SAEs were considered mild or moderate for over half of patients who reported SAEs within each group. The incidence of severe events was also similar in both groups. A total of 137 deaths occurred during the study, 65 patients in the NB32 group and 72 in the placebo group, although no deaths in this study were related to the study drug.

Table 50: Overall summary of TEAEs leading to discontinuation, TESAEs and all deaths, totality of data

	NB32 (n=4455)	Placebo (n=4450)
Drug-related TEAEs, n (%)	982 (22.0)	174 (3.9)
Severe TEAEs, n (%)	217 (4.9)	108 (2.4)
TESAEs, n (%)	463 (10.4)	386 (8.7)
DC due to AEs, n (%)	1292 (29.0)	400 (9.0)
Deaths, n (%)	65	72
Key: AE, adverse event; DC, discontinuation; TEAE, treatment emergent adverse event; TESAE, treatment emergent serious adverse event. **Source:** Orexigen, 2015[92]		

As shown in Table 51, discontinuations due to AEs most commonly included GI AEs, which occurred in 14.2% of NB32 patients and 1.9% of placebo-treated patients (p<0.001), and central nervous system symptoms, which occurred in 5.1% 1.2% of patients, respectively (p<0.001). Psychiatric symptoms resulted in study drug discontinuation in 3.1% of NB32 patients and 0.9% of placebo patients (p<0.001).

Table 51: Most common adverse events leading to discontinuation of study drug, NB-CVOT study

Adverse event, n (%)	NB32 (n=4455)	Placebo (n=4450)
Any AE	1292 (29.0)	400 (9.0)
Gastrointestinal	631 (14.2)	84 (1.9)
Nausea	333 (7.5)	21 (0.5)
Constipation	123 (2.8)	15 (0.3)
Vomiting	87 (2.0)	1 (<0.1)
Central nervous system	226 (5.1)	51 (1.2)
Tremor	77 (1.7)	0
Dizziness	62 (1.4)	7 (0.2)
Headache	51 (1.1)	14 (0.3)
Psychiatric disorders	136 (3.1)	39 (0.9)
Insomnia	35 (0.8)	16 (0.4)
Anxiety	26 (0.6)	8 (0.2)
Hallucinations	11 (0.2)	0
Depression	5 (0.1)	9 (0.2)
Increased blood pressure	39 (0.9)	23 (0.5)

Adverse event, n (%)	NB32 (n=4455)	Placebo (n=4450)
Palpitations	19 (0.4)	5 (0.1)
Feeling jittery	15 (0.3)	1 (<0.1)
Flushing or hot flashes	13 (0.3)	2 (<0.1)
Fatigue	12 (0.3)	1 (<0.1)
Key: AE, adverse event; NB32, naltrexone 32mg plus bupropion **Source:** Nissen et al. 2016[52]; Orexigen, 2015[92]		

The NB-CVOT trial provides supportive safety data from a large population of almost 9,000 patients, which demonstrates that, despite the higher risk patient population, NB32 was well tolerated, even in patients receiving anti-depressants. Furthermore, it is important to note that regulatory bodies approved NB32 on the basis of the CV risk, or lack thereof, demonstrated within this study.

IGNITE study

In the IGNITE study, the safety profile shown was consistent with that seen in the previous, pivotal trials; most patients tolerated NB32 well, and those who developed AEs did so early in the treatment protocol.[54] The most common AE leading to NB discontinuation was nausea (7.0% of all subjects), which is consistent with the rate in the Phase III trials (6.3%). Only two AEs led to discontinuation in the NB32 plus standard management group after Week 26 (both with AE onset before to Week 26), and no AEs necessitating discontinuation had an onset date during the extended time period (Weeks 52-78).

Comparative safety

Throughout the four COR trials, NB32 was shown to have a tolerable safety profile in line with that previously seen in trials of the two component drugs. A previous systematic review and meta-analysis was conducted which suggested that NB32 is not as well tolerated as the current pharmacological treatment option, orlistat, based on discontinuation due to AE data (OR: 1.44 [95% CI: 1.07, 1.95]).[93] However, it is important to note that, due to the different mechanisms of action, the safety profiles of each drug are very distinct; indeed, the AEs that accounted for discontinuation in the orlistat trials are generally considered to be much more debilitating than those seen in the trials of NB32.

In the four key NB32 trials, along with a Phase II study, the most common reason for discontinuation was nausea, seen in 6.3% of patients with most of this withdrawal occurring during the dose-escalation phase.[53] Nausea was also one of the most commonly reported (≥5% in either group) side effects, reported in 31.1% of non-diabetic and 42.3% of diabetic patients.[53] Other frequent AEs with NB32 treatment include constipation, vomiting, dizziness, dry mouth, headache and insomnia; all are side effects consistent with the AE profiles for the individual drug components.

Discontinuations due to individual events other than nausea were <2% in NB32 treated patients.[53] Furthermore, the majority of discontinuations due to AEs occurred in the dose-escalation phase (17.4%) with only 23.8% of patients discontinuing due to AEs across the double-blind treatment period.[53] This suggests that events such as nausea are more pronounced at the beginning of treatment. Nausea peaked within 4 weeks and resolved in most patients by 24 weeks.[53] In addition, no events of nausea were considered serious, and it should be noted that anti-nausea medication, although appropriate for use in patients receiving NB32, had very limited use within the trials.[53]

By contrast, the most frequently reported side effects seen with orlistat mainly consist of disabling and incapacitating GI effects such as oily spotting from the rectum, flatus with discharge, fatty or oily stools and increased defecation. GI side effects accounted for almost half (49.4%) of all spontaneous AEs seen in the orlistat trials.[28] Such effects can severely limit daily activity as patients need to carefully consider the proximity of toilets and changing rooms each time they leave home. These considerations, alongside embarrassment caused by such conditions, can impair patient's social life as they are less likely to continue with normal social activities. This in turn can lead to patients becoming isolated.

Orlistat is also associated with risks of liver reactions, and the EPAR report states that treatment with orlistat can result in hepatitis, which may be serious, and increases in transaminases and alkaline phosphatases.[28] Out of a total of 846 hepatic events seen across trials with orlistat, 271 were serious and included a total of 21 cases of serious liver toxicity where the role of orlistat cannot be definitively excluded. This included five cases of hepatic failure, which led to death in two cases and liver transplantation in three cases.[28] These increased risks are all addressed within the risk management plan and, although analysis of spontaneous reports

suggest only weak evidence of a causal relationship, such causality between orlistat and hepatic events cannot be excluded. By contrast, across the pivotal NB trials described above, only one death occurred due to a CV event, and this was judged to be unrelated to treatment with NB32.[53]

In NB32 treated patients, AEs relating to liver toxicity were seen in only 1.2% of patients and were mostly due to elevated transaminases; only 0.2% of NB32-treated patients discontinued treatment due to elevated liver enzymes and there were no cases of hepatic failure observed in the Phase 3 studies.[53]

4.13 Interpretation of clinical effectiveness and safety evidence

There is an increasing prevalence of adults who are overweight or obese with one or more weight-related comorbidities. These conditions are associated with a large and increasing patient, caregiver and economic burden. Despite this high burden, only one pharmacological treatment, orlistat, is currently available for overweight and obese patients; however, it is associated with serious debilitating side effects and waning effectiveness (see Section 3.6).

There is a clear unmet medical need for additional, effective and well-tolerated pharmacological therapies that can induce and sustain weight loss in patients who have not achieved adequate weight loss through dietary and exercise changes.

Main findings from clinical evidence base

The clinical benefits and potential harms associated with NB32 have been demonstrated with clinical data from four pivotal Phase III RCTs alongside two longer-term Phase IIIb trials. Principal findings from this evidence base are summarised below:

<u>Early and sustained weight loss that was significantly greater than that observed with standard management without NB</u>

In all four pivotal trials, weight loss began as early as Week 4 in NB32-treated patients, and this continued across the duration of the 56-week trials. Furthermore, percentage weight loss at 56 weeks and the proportion of patients who lost ≥5% body weight at Week 56 was significantly greater ($p<0.001$) for patients treated with NB32 compared to patients who received standard management (the placebo group). It is important to note that such clinical benefit was observed across all

patient groups, including those with obesity and those who were overweight in the presence of one or more weight-related comorbidities, including patients with T2DM.

Across the four pivotal trials, the proportion of patients with ≥5% weight loss at Week 16 ranged from 44.9% to 69.9%; however, because the 16-week discontinuation rule was not a feature of these trials, all patients continued in the study. As such, the reductions in weight seen at Week 56 should be viewed as a conservative estimate given that in clinical practice approximately half of these patients would have discontinued study treatment. The economic modelling accounts for this by using only data from responding patients (see Section 5). In pooled analysis of patients who were Week 16 Responders and continued to receive treatment up to 56 weeks, the LS mean weight loss was 11.7%, with 57% of these patients losing ≥10% of their original bodyweight. Such a reduction in bodyweight is well accepted to improve overall health and reduce the risk of developing weight-related complications.[55]

Importantly, two Phase IIIb RCTs showed weight loss was sustained across longer term treatment with NB32, including in patients with more severe cardiovascular risk factors, as in the case of the NB-CVOT study.

Significant improvements in many cardiometabolic parameters and diabetic-specific risk factors

Alongside significant reductions in weight, NB32 was associated with significant improvements in numerous cardiometabolic parameters, which could potentially lead to a reduction in the risk of CV events.

In ‘pre-diabetic’ patients (COR-I, COR-II, COR-BMOD), fasting insulin and HOMA-IR (a measure of insulin resistance) levels were significantly reduced when treated with NB32 compared to placebo ($p<0.005$ for both outcomes) across three of the pivotal studies, which could potentially reduce the risk of these patients to develop T2DM. In patients with T2DM (COR-DM), levels of HbA1c were significantly reduced after treatment with NB32 and a greater proportion of patients achieved standard treatment targets of HbA1c levels ($p<0.01$ vs placebo), showing that NB32 can improve glycaemic control.

Significant and sustained improvements in disease-specific HRQL, as measured by the IWQOL-Lite tool

Significant improvements in the IWQOL-Lite total score was seen across all three pivotal trials in 'pre-diabetic' patients treated with NB32 ($p<0.05$ vs placebo). A numerically greater improvement was also seen for diabetic patients treated with NB32, compared to placebo, in the COR-DM study. Although, this did not reach significance, this likely reflects the greater disease burden associated with this diabetic population. Nonetheless, it is encouraging to see an improvement in weight-related quality of life, as with longer-term reductions in weight, the burden of T2DM may also be alleviated leading to more significant improvements in HRQL. Improvements were also seen in the physical function and self-esteem subscales ($p<0.01$ for NB32 vs placebo in the COR-I and COR-II studies); these improvements were maintained to Week 56 across the pivotal studies.

Significant and sustained improvements in control of eating and reduced food cravings, as measured by the COE questionnaire

NB32 treatment resulted in improvements in the COE questionnaire, indicating reduced hunger and strength of food cravings, as well as increased feelings of fullness and ability to resist food cravings ($p<0.05$ for all comparisons across studies). These improvements generally persisted for the duration of the trial and, importantly, were seen in the diabetic as well as the 'pre-diabetic' population.

Reduced food craving could be related to the innovative mechanism of action seen with NB32. NB32 targets hypothalamic regions responsible for appetite and energy expenditure, and which in humans is thought to lead to reduced hunger. Furthermore, NB32 targets mesolimbic reward circuits, which influence reward pathways for eating behaviours, thus modulating food craving and mood.[25]

ITC demonstrates at least comparable efficacy for NB32 compared with current drug management, orlistat

Results from the ITC (presented in Section 4.10) suggest that NB32 is more efficacious than placebo and at least as efficacious as orlistat. In patients, without T2DM, both placebo and orlistat are seen to have statistically significant inferiority versus NB32 for 5% reduction in weight ($OR_{placebo}$: 0.24 [95% CrI: 0.20, 0.29] and $OR_{orlistat}$: 0.77 [95% CrI: 0.61, 0.96]) and for mean percentage weight change from

baseline ($MD_{placebo}$: 4.88 [95% CrI: 4.35, 5.43] and $MD_{orlistat}$: 1.13 [95% CrI: 0.44, 1.80]) at 1 year.

In patients with T2DM, weight loss may be more difficult and therefore any weight loss is seen to be beneficial. The ITC results suggest that NB32 is more efficacious than placebo, which is seen to have statistically significant inferiority versus NB32 for 5% reduction in weight (OR: 0.29 [95%CrI: 0.18, 0.46]) and for mean percentage weight change from baseline (MD: 3.21 [95% CrI: 2.23, 4.21]) at 1 year. NB32 is seen to have comparable efficacy with orlistat, with neither treatment showing statistical superiority for either outcome.

Sensitivity analyses were also performed to explore heterogeneity between studies by excluding subgroups of studies (presented in Section 4.10). The results of these analyses produced relative treatment effect estimates which were consistent with the base case analyses.

<u>NB32 is generally well tolerated, with a transient and manageable AE profile</u>

Across the four pivotal trials, NB32 was well tolerated, with a transient and manageable AE profile that clinicians will be familiar with due to use of the individual components of the drug, despite the differing doses used. Most TEAEs were considered mild or moderate in severity, with severe TEAEs <20% across all studies. In addition, few TESAEs were observed across all trials, with rates of <4% across all pivotal studies and as low as 1.6% in the COR-I study. Across the pivotal studies, there was only 1 death in patients treated with NB32 (n=2,482), which was considered unlikely to be related to the study drug.

Common TEAEs were transient and, in most cases, did not lead to discontinuation. Nausea was the most common AE leading to discontinuation and was reported in 6.3% of patients.

All NB32 trials showed a consistent safety profile, observed across all patient groups including those with obesity and those who are overweight in the presence of one or more weight-related comorbidities, including patients with T2DM with no difference in rates of hypoglycaemia between treatment groups. Compared to orlistat, the only pharmacological treatment option currently available, NB32 offers a less disabling and incapacitating safety profile, which in turn could better allow patients to maintain

their quality of life, and remain on treatment for a longer period. This could further contribute to more meaningful and sustained weight loss.

The longer-term safety of NB32 is supported by the IGNITE study, and the NB-CVOT study which demonstrates that NB32 is well tolerated even in an older patient population with CV disease, many of whom also had T2DM, hypertension and/or received anti-depressant therapy. Indeed, regulatory approval was granted on the basis of this study showing a lack of MACE and CV risk.

Strengths and limitations of clinical evidence base

Overall, the clinical evidence provides an appropriate base to inform the assessment of clinical and cost effectiveness of NB32 for the management of weight in adults who are obese, or overweight with one or more weight-related comorbidities.

The clinical effectiveness of NB32 was assessed across a large clinical trial programme including 25 completed trials. This provided RCT evidence from four pivotal Phase III studies, and longer-term evidence from the NB-CVOT and IGNITE studies. All four pivotal trials, were conducted in line with GCP guidelines, with steps taken to minimise bias and independent monitoring or advisory committees in place to provide oversight of safety and efficacy considerations, study conduct and risk-benefit ratio. These trials provided evidence for a wide range of patient groups that are representative of patients who would present in clinical practice after failing to achieve adequate weight loss on standard management. Furthermore, although no UK centres were included in the four pivotal trials, clinician feedback confirmed that the patient population included in the trials was a fair reflection of the average patient seen in UK NHS practice, although the mean BMI of 36kg/m^2 was slightly higher than usually seen in clinical trials.[31] In addition, consistently superior clinical benefit was observed with NB32 compared to standard management across all pre-determined subgroups, including those with hypertension and dyslipidaemia and in patients with T2DM.

All four pivotal trials directly compare NB32 to placebo; however, patients in both arms received some form of standard management, consisting of diet instruction, advice on behaviour modification and physical activity suggestions. As such, the placebo arms in all trials reflect the standard management patients would receive in the absence of NB32 being available; one of the named comparators in the decision

problem. NB32 should therefore be considered an alternative first-line pharmacological treatment in patients for whom orlistat is contraindicated or is not utilised due to physician/patient choice, and who currently persevere with standard management in current practice, despite the expected lack of effectiveness.

Currently, orlistat, the other named comparator in the decision problem, is the only weight control medicine widely available in the EU. Active comparator trials of NB32 compared to orlistat were not considered appropriate as the distinct tolerability profile of orlistat makes it difficult to blind.[28] Using orlistat as an active reference could have led to un-blinding of patient treatment allocation and potentially to disparate patient withdrawal patterns. As such, no NB32 trials included an active reference, and this omission was deemed acceptable by the CHMP.[53] Although head-to-head data are not available for orlistat, an NMA has been conducted, which demonstrates numerically greater benefit of NB32 compared to orlistat with regards to reduction in body weight. NB32 was also shown to be statistically superior to standard management in all outcomes (see Section 4.9.1).

NB32 has not been investigated in patients who have not achieved adequate weight loss with orlistat; however, clinician feedback has confirmed that having previously received orlistat is not expected to have any effect on the efficacy of NB32.[31] This is primarily due to the distinctly different mechanisms of action of the two drugs. As such, NB32 offers a pharmacological treatment option to patients who have not achieved adequate weight loss with orlistat treatment, or who did not comply with dietary requirements associated with orlistat, or were unable to tolerate orlistat treatment and would otherwise revisit standard management measures.

The four pivotal studies employed primary endpoints of mean and categorical changes from baseline in body weight, as well as various secondary endpoints. In aggregate, these efficacy endpoints allowed for a thorough investigation of the effect of NB32 on weight loss/maintenance, cardiometabolic and diabetic risk factors, HRQL and eating behaviour. In addition, the COR-DM study permitted an assessment of the efficacy of NB32 on glycaemic control in patients with diabetes.

Based on CHMP recommendations, the primary demonstration of efficacy should be based on a difference in mean weight loss from baseline of ≥10%, and ≥5% in the active treatment group compared to placebo.[55] Although the trials' primary efficacy

endpoints were based on a 5% criterion to meet US guidance, each trial also included a prospectively defined 10% categorical weight loss secondary endpoint. Therefore, the clinical efficacy programme allowed a comparison of the responder rate in terms of the proportion of subjects who met the more stringent categorical weight loss criterion of ≥10% weight loss. This was deemed acceptable to the CHMP.[53] There were also concerns of potential bias in the primary efficacy analysis highlighted by the CHMP. However, pre-defined sensitivity analyses supported the primary efficacy analysis outcomes such that these concerns were not substantiated, and marketing authorisation was approved with an agreement to present more conservative data in the SmPC.

In summary, the clinical evidence shows that NB32 addresses the clear unmet need seen within overweight and obese patients. NB32 offers an alternative pharmacological treatment option with an improved, multi-modal mechanism of action and favourable AE profile to patients treated with orlistat in current clinical practice. Perhaps more importantly, NB32 offers a pharmacological treatment option with proven weight loss efficacy to patients treated with standard management in the absence of a better treatment option in current clinical practice.

4.14 Ongoing studies

There are no ongoing studies that are anticipated to provide data of relevance to the decision problem within the next 12 months.

5 Cost effectiveness

5.1 Published cost-effectiveness studies

A systematic search was performed to identify published economic modelling studies evaluating pharmacological treatments for obese individuals, or overweight individuals with one or more comorbidity. The targeted databases were MEDLINE, MEDLINE in process, Embase, Cochrane Library, NHS EED and CRD HTA.

A detailed search strategy is provided in Appendix 15.

The relevance of each reference for data extraction was assessed based on pre-specified eligibility criteria. The criteria used are summarised in Table 52.

Table 52: Eligibility criteria for economic modelling evidence search

Criteria	Inclusion	Exclusion
Population	• Adults who are obese (BMI ≥30kg/m^2), or overweight (BMI ≥25kg/m^2, adopting the most inclusive criterion from the summary of product characteristics and care guidelines: that used in NICE Clinical Guideline 189) with one or more comorbidities (T2DM, dyslipidaemia and/or controlled hypertension)	• Healthy volunteers • Children (age <18 years) • Diseases other than those specified in inclusion criteria
Intervention/ comparator	• At least one pharmacological or weight management intervention for obesity assessed in the model	• Studies were not excluded based on comparator therapy
Outcomes	• ICER • Costs (unit and total) • QALYs • LYs • Incremental costs • Incremental QALYs/LYs • Model inputs (e.g. transition probabilities) • Sensitivity analyses results	
Study type	Full economic evaluations, such as: • Cost-consequence • Cost-effectiveness • Cost-utility • Cost-benefit (Cost-minimisation, cost-saving and budget impact analyses were included at the secondary screening stage but data from these	• Non-systematic reviews[a] • Letters • Comment articles • Burden of illness studies • Non-modelling studies

	studies were not extracted with other modelling studies[b]. Relevant cost and resource use data for UK population from these studies, were extracted with other cost and resource use studies.)	
Language	• Studies published in English • Studies published in languages other than English[c]	
Publication timeframe	• Studies published in or after 2006 (last 10 years)	• Published before 2006

Key: BMI, body mass index; ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life year; T2DM, Type 2 diabetes mellitus.
Note: [a], Systematic reviews were included and flagged for bibliography searches; [b], Cost-saving, cost-minimisation and cost of illness studies and budget impact analyses that did not report any cost-effectiveness data or did not relate cost to outcomes were not extracted; [c], Studies published in languages other than English would be explored only if insufficient evidence found.

The PRISMA diagram in Appendix 15 presents the flow diagram of studies identified for the cost-effectiveness review. In total, 1,781 citations were identified through database searching, with one additional citation identified through bibliographic searching and 10 abstracts identified from conference proceedings.

Following screening and eligibility assessment, 22 publications were identified from which a total of 19 studies were included in the review.[14, 94-111] Tabular summaries of study characteristics and results are provided in Appendix 15.

The studies included in the review varied in terms of model type, geographical location and pharmacological intervention(s) considered. None considered NB32 as an intervention. Four identified studies were set in the UK. One of these was a 2012 *Health Technology Appraisal* report published by Ara et al. comparing different pharmacological treatments for obesity[14], another was a critique of the manufacturers submission to NICE for rimonabant[108], and two were cost-utility analyses in patients with T2DM.[107, 109]

Evidence from the review suggests that pharmacological treatment for obesity has the potential to be highly cost effective. Uncertainty analyses in previous studies showed results to be particularly sensitive to uncertainty surrounding assumptions concerning duration of weight maintenance after initial weight loss and assumptions around the effect of a reduction in body mass index (BMI) on HRQL.

Following completion of the systematic search, in July 2016, NICE and Public Health England (PHE) websites were searched for any further evidence of interest. The search of the NICE website returned no additional TA documents of interest; NICE Clinical Guideline (CG) 43 was published in 2010[47], superseding recommendations from TA22 (Obesity – orlistat) and TA31 (Obesity – sibutramine). In 2014, NICE CG 189 was published as an update to CG 43, but a cost-effectiveness appraisal of pharmacological weight management treatment was not scoped within this update.[3] The search of the PHE website did however identify their "Weight Management Economic Assessment Tool": a model designed to help healthcare professionals assess existing or planned weight management interventions and to allow commissioners to compare the costs of an intervention for English patients with potential cost savings.[112]

Overall, previous economic analyses have varied in terms of their usefulness to inform the decision problem. A variety of model types and structures have been used across studies, with most studies using timed cohort models. However, to capture the cost and health consequences of weight reduction strategies, it is important to capture consequences for both weight and weight-related events with chronic implications. When modelling events with chronic implications, keeping track of patient histories has great importance. In a cohort model, this can be achieved by creating health states to differentiate between patients with different histories. The number of health states required for this can quickly become very large, and difficult to manage. This has been a limiting factor in much of the identified evidence base.

Only one study in the review, that reported by Ara et al.[14], used an individual-level timed model and avoided the inherent limitations of cohort models in this disease area. This study has numerous additional advantages as a source of evidence for this submission. The authors used a systematic approach to search, appraise and synthesise evidence, with the stated aim of evaluating the clinical and cost-effectiveness of pharmacological treatments for overweight or obese patients, from the perspective of the UK NHS and Personal Social Services.[14] Thus, Ara et al. developed epidemiological models of how changes in BMI affect the risk of major clinical events, and how BMI levels change as a population ages, through analysis of longitudinal data from the General Practice Research Database (GPRD).[14] Outputs from these natural history models, plus analyses of the relationship between EQ-5D

utility data and BMI using individual-level Health Survey for England (HSE) data also disseminated within the report, are particularly relevant for a novel appraisal in this field.[14] In addition, the space for reporting afforded in the publishing journal, Health Technology Assessment, is far greater than that of many journals, with the criterion for "systematic" in the Health Technology Assessment journal series being the theoretical permission of replication of the review by others.[14]

5.2 De novo analysis

5.2.1 Patient population

As described in Section 2.2, NB32 is licensed as an adjunct to standard non-pharmacological management in adult patients who are obese (BMI ≥30kg/m^2) or overweight (BMI ≥27kg/m^2 and <30kg/m^2) in the presence of one or more weight-related comorbidities.[29] The *de novo* economic analysis evaluates the cost effectiveness of NB32 in this patient group.

The key clinical data for the economic analysis are from the four multicentre, randomised, double-blinded, placebo-controlled studies comprising the COR trial programme (COR-I, COR-II, COR-BMOD and COR-DM), described in detail in Section 4.7. Participants in three of these studies (COR-I, COR-II, COR-BMOD) were adults with BMI 30–45kg/m^2 or BMI 27–45kg/m^2 and dyslipidaemia or controlled hypertension. Participants in the COR-DM study were adults with T2DM and BMI 27–45kg/m^2. Patient characteristics from COR trials were validated as reflective of the typical patient group who would stand to benefit from NB32 in UK NHS practice, at clinical review.[31]

Comparisons to orlistat adjunct therapy and standard management alone, for the non-T2DM and T2DM patient groups in the COR trial programme, are supported by results from NMAs, described in Section 4.10.

A key limitation of COR trials and other trials in the meta-analyses in Section 4.10, for the purposes of a lifetime economic evaluation, is length of follow-up. The IGNITE study[54], described alongside COR trial programme studies in Section 4.7, followed patients for 78 weeks, but patient numbers beyond 52 weeks were low; 61 NB32 patients were followed from Week 52 onwards. A far greater number of NB32 patients were followed beyond 52 weeks in the NB-CVOT study[52] (n=748; Figure 31,

Section 5.3.2), as also described in Section 4.7. NB-CVOT study data are used to inform assumptions about treatment continuation and effectiveness beyond 1 year in the economic analysis. Although BMI inclusion criteria were only slightly different to those for COR trials (BMI ≥27kg/m^2 and <50kg/m^2), patients in the NB-CVOT study were older than those in the COR trial programme (with inclusion restricted to men over 45 years and women over 50 years), and enrolment was restricted to patients with increased risk of cardiovascular outcomes.[52, 92] Nevertheless, evidence from this study is valuable in informing assumptions beyond short-term trial endpoints.

The natural history models developed by Ara et al. of lifetime BMI and risks for the development of key weight-related disease and death, are pivotal for the economic appraisal of NB32.[14] As described in Section 5.1, these models are based on patient data from the GPRD*. Ara et al. accessed the GPRD in January 2011.[14] At the time of access, the GPRD contained anonymised primary care records from over 12 million patients in the UK. Ara et al. drew a sample of longitudinal patient records from adult patients who had three or more BMI readings of over 27kg/m^2, as a basis for their BMI risk model and BMI natural history model analyses.[14]

5.2.2 Model structure

5.2.2.1 Model type

A *de novo* individual-level economic model was developed for this appraisal. As described in Section 5.1, an individual-level approach is better suited than a cohort-level approach to capture the chronic implications of both weight and weight-related health events in a heterogenous group of overweight and obese patients.

The *de novo* economic model harnesses many assumptions and key input data from Ara et al.[14], including the natural history models for BMI lifetime patterns and BMI risks estimated using GPRD data. These statistical models were specified in synthesis with the economic model developed by Ara et al.; their outputs are suited to inform an individual-level approach. Specifying a *de novo* model that can make

* In March 2012, the GPRD became part of the Clinical Practice Research Datalink.113. Medicines & Healthcare products Regulatory Agency. The Clinical Practice Research Datalink. 2016. Available at: https://www.cprd.com/home/. Accessed: 10 November 2016.

best use of the input data available is another key reason for selecting an individual-level modelling approach. For the same reason, the model treats time as continuous, which is a natural selection given the underlying data.

5.2.2.2 Model software and DICE methodology

Ara et al. built their model in Simul8® software. This software is well-suited to individual-level economic modelling. However, its use is problematic for the purposes of HTA in many jurisdictions, owing to the financial cost of a license and the need for the technical review team to be comfortable with the specialist software.

Microsoft Excel® has the advantage over specialist software such as Simul8® of being familiar and transparent to HTA review teams and stakeholders, and is sufficiently flexible for the specification of the *de novo* model. However, building an individual-level economic model in Excel necessitates a greater reliance on underlying Visual Basic for Applications® (VBA) code than is typical for cohort-level economic models. If this code is well annotated and set out, transparency should persist for reviewers familiar with VBA language, but may otherwise suffer. Caro recently proposed a "discretely integrated condition event" (DICE) approach to structure a pharmacoeconomic decision problem as a set of conditions (aspects that persist over time) and events (aspects that occur at a point in time) within spreadsheet tables that specify condition values and event consequences.[114] Although not distinct from a typical approach to individual-level modelling, the principles and structure outlined by Caro were used in the *de novo* model to maximise transparency and clarity.[114]

The DICE approach is suggested to present a unifying approach that has been deliberately designed to meet the modelling requirements in a straightforward transparent way, without forcing assumptions (e.g. only one transition per time cycle) or unnecessary complexity.[114] Detailed explanation of the DICE approach is disseminated by Caro.[114] However, a top-level overview of the approach is provided here to aid interpretation.

Within the DICE approach, a disease and its management are conceptualised in terms of "conditions" and "events". A condition is something that persists or occurs over a period of time[114], and is characterised by a name and label. An example condition in the *de novo* model is named "di_BMI". This condition describes the

simulated patient's BMI, and its label takes the value of a positive real number. Another condition in the *de novo* model is named "di_female". This condition describes the simulated patient's gender, and its label can take the value of 1 if the patient is female and 0 otherwise. A third example is the condition "di_diabetic_status", which has the value 1 if the simulated patient either enters the model with T2DM or develops T2DM during the modelled time horizon. The use of conditions allows the history and status of simulated patients to be tracked easily.

An event is distinct from a condition and occurs at a point, or points, in time. An event may initiate or terminate a condition, modify its level or affect the occurrence of other events.[114] Each event has an associated time of occurrence. An example event in the *de novo* model is "di_diabetic_onset", to capture the event of the onset of T2DM. The time of occurrence associated with "di_diabetic_onset" is estimated in the model for each patient, using the BMI risk equation for T2DM described in Section 5.3, for all patients who do not enter the model with T2DM. The event "di_diabetic_onset" triggers a change in the value of condition "di_diabetic_status" from 0 to 1. As described in Sections 5.3, 5.4 and 5.5, diabetic status affects mortality risk, HRQL and healthcare costs in the *de novo* model.

5.2.2.3 Model structure and logic

An overview of the *de novo* model structure is shown in Figure 25. One simulated individual (defined by their sampled baseline characteristics) is followed to death three times before the next enters the model. In the first patient run, the patient is assigned NB32 as adjunct therapy; in the second patient run, the patient is assigned orlistat as adjunct therapy; and in the third patient run, the patient is assigned standard management only. Random numbers are assigned for a patient, and therefore, across treatment arms, the same random numbers are utilised. This approach was helpful for validation and verification.

The following narrative, alongside Figure 25, aims to describe the logic and assumptions underpinning the model, by describing a simulated patient's journey through the model.

Figure 25 depicts the process of a simulated individual's progress through the model, from model entry ("START"), through the various treatment and disease events that

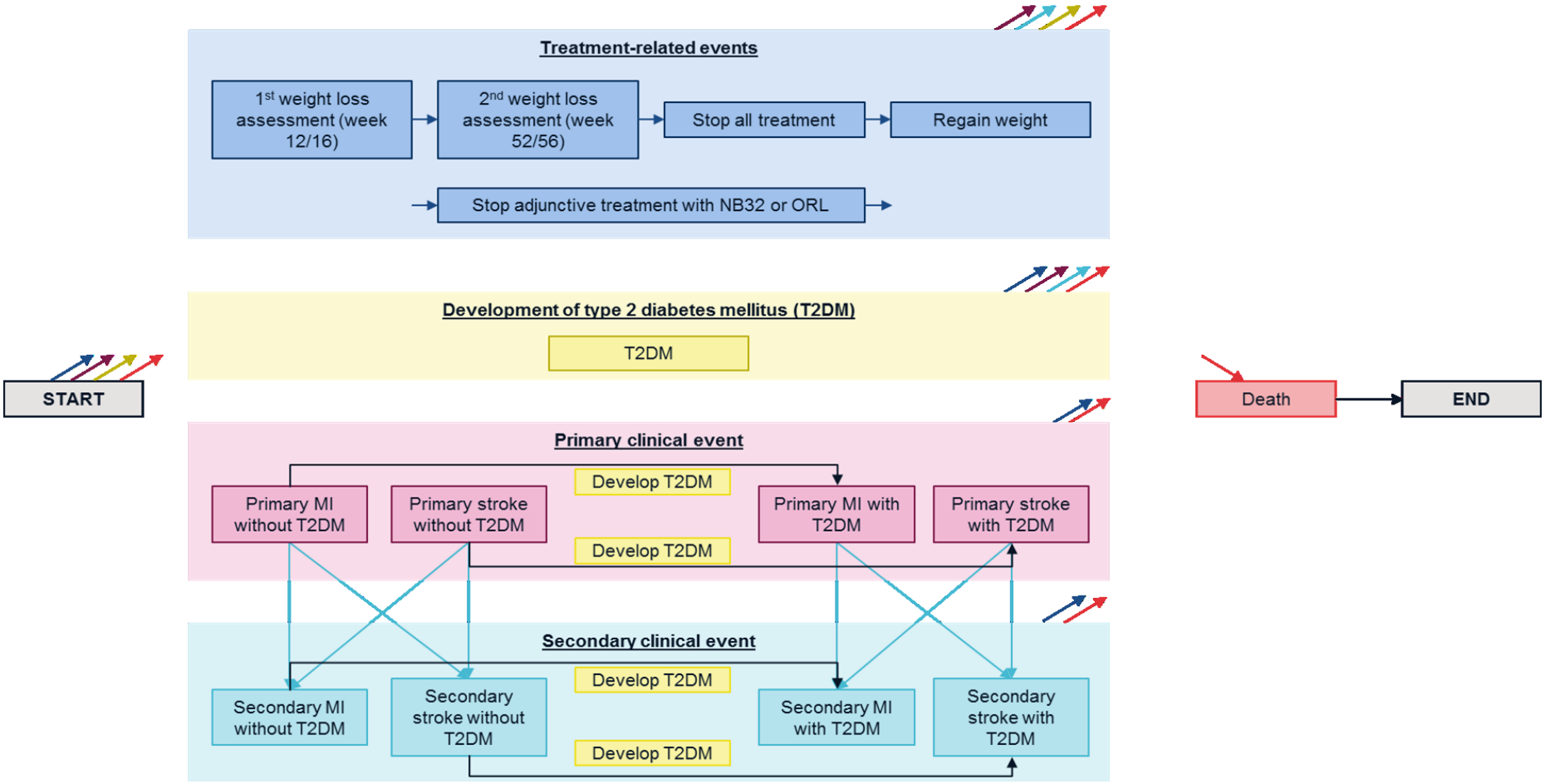
may occur in the model and have consequence for patient utility and/or health and social care costs, to death and model exit (“END”).

Upon model start and patient entry

Upon model entry, time is zero, and a simulated patient is assigned a baseline profile of characteristics that are explanatory factors for risks, costs or utility in the model, that are stored as conditions. The baseline profile characterises the individual by gender, age, T2DM status, other type diabetes mellitus status, BMI, height, binary categories to capture use of aspirin, insulin, statins and blood pressure treatment, whether the individual is an ex-smoker, and whether the individual is a smoker. The sources for these data are presented in Section 5.3.1. Following attribution of baseline characteristics, the individual is attributed the condition for NB32 adjunct therapy.

Then, the individual is assigned sampled “time to event” (TTE) values for each event they are at risk of experiencing next. As illustrated through colour coding in Figure 25, this could be a treatment-related event or a disease-incidence event.

Figure 25: *De novo* model diagram



Key: MI, myocardial infarction, NB32, naltrexone 32mg plus bupropion; ORL, orlistat; T2DM, Type 2 diabetes mellitus.
Notes: Arrows demonstrate the possible transitions to each type of event.

As depicted in Figure 25, some treatment events have fixed times. As this is the first patient run, response rate assessments for adjunct NB32 therapy are set as events at 16 and 56 weeks.† It is possible for the simulated patient to discontinue adjunct NB32 therapy before, between or after these timepoints, but in line with licence wording, if a simulated patient receiving adjunct NB32 therapy has not discontinued adjunct therapy before 16 weeks, and their estimated weight loss from time zero is <5%, NB32 is discontinued. In line with clinical guidance from John Wilding, Professor of Medicine and Consultant Physician with extensive experience of treating overweightness and obesity in NHS patients[31], a similar rule is applied at 56 weeks (12 months after the end of the 4-week titration period).

Other events do not have fixed times. These are times to: discontinuation of adjunct therapy, discontinuation of standard management, disease incidence, death; they can occur before, between or after the scheduled response assessment events. Times to these events are estimated and recorded for the patients using the data described in Section 5.3.

First event

When times to each event have been estimated, a condition for current time is updated to the time of the first event.

Next, any conditions affected by the first event are updated. For example, if the patient is predicted to stop adjunct NB32 treatment before the first scheduled response assessment, a condition is used to record that the individual is no longer receiving adjunct treatment.

Following updating of conditions, TTE estimates are updated for any events affected by condition changes from the first event. For example, if an event changes BMI, times to obesity-related-disease events (for which BMI variables have explanatory power [Section 5.3]) are re-estimated. Revaluation of TTE estimates are calculated within the model using Equation 1. This equation considers the originally sampled TTE and recalculates this based on conditions that have changed since this time (e.g. an increase in BMI).

† 12 and 52 weeks for the second patient run, where adjunct orlistat is assigned

Equation 1: Calculating change in event time given initially sampled time

$$TTE_{recalculated} = Time_{current} + TTE_{re-sample}\left(\frac{TTE_{original} - Time_{current}}{TTE_{original}}\right)$$

Key: $Time_{current}$, current time; $TTE_{recalculated}$, Recalculated time to event based on patient being event-free until the current time; $TTE_{original}$, Originally sampled time to event; $TTE_{re-sample}$, Re-sampled time to event from time = zero (i.e. does not consider the current time).

In Equation 1, $Time_{current}$ is the current time in years. $TTE_{original}$ is the originally sampled TTE estimate, taken from the previous event. $TTE_{re-sample}$ is the re-sampled TTE estimate, based on conditions that have changed since the calculation of $TTE_{original}$ (e.g. a change in BMI). $TTE_{recalculated}$ is then derived using these estimates, and becomes $TTE_{original}$ should the TTE estimate in the next event need to be re-calculated.

The total costs and QALYs accrued up from time zero to current time are then calculated, using utility and cost assumptions described in Sections 5.4 and 5.5, and recorded as further conditions that act as running totals for the individual.

Discounted total costs and QALYs are also calculated, by discounting the life years accrued between model entry and first event. This is done by calculating the integral of the exponential survival curve, between last event and current time. The formulae for this calculation are shown in Equation 2:

Equation 2: Discounting life years

$$LYs_{discounted} = \frac{e^{(-DR_{inst} * Time_{current})}}{-DR_{inst}} - \frac{e^{(-DR_{inst} * Time_{previous})}}{-DR_{inst}}$$

where $DR_{inst} = \ln(1 + DR)$

Key: DR, annual discount rate; DR_{inst}, instantaneous discount rate; $LYs_{discounted}$, discounted life years; $Time_{current}$, current event time; $Time_{previous}$, previous event time.

In Equation 2, DR is the annual discount rate for costs and health outcomes, specified as 3.5% per annum in the NICE reference case.[115] DR_{inst} is the instantaneous discount rate, and $Time_{current}$ and $Time_{previous}$ are measured in years. Discounted total costs and QALYs are recorded by further "running total" conditions.

Subsequent events

Following the first event, subsequent events occur in chronological order for the patient. The processes described for the first event, which are to (i) update time, (ii) update conditions, (iii) recalculate any time-to-event estimates affected by condition changes and (iv) calculate and record updated total and discounted total patient costs and QALYs, are repeated for each subsequent event, until the next event is death.

Upon death and model end

Upon death, the individual's lifetime total costs and QALYs are documented in the model. Time is reset to zero, treatment conditions are reset, and "Upon model entry" logic begins the process for the next patient run.

If the next patient run is for orlistat adjunct therapy or for standard management alone, the random number profile will not change, and the patient characteristics will remain. If the next patient run is for NB32 adjunct therapy, a different random number profile is selected. This process continues until the selected number of patients have been run through the model.

5.2.2.4 Methodological modelling assumptions

Many factors contribute to obesity, and they relate to each other in non-linear fashions, are subject to time delays, and change over time.[116] Inherently, to appropriately model a complex condition such as obesity, a number of modelling assumptions must be made.

Table 53 shows the key methodological modelling assumptions utilised within the *de novo* economic model constructed to inform this submission of evidence. These assumptions relate to the modelling of obesity in general, as opposed to assumptions relating to available data.

Table 53: Key modelling assumptions utilised in the *de novo* economic model

Assumption made	Rationale
Treatment discontinuation	
If a patient discontinues treatment with NB32 or orlistat, it is assumed that the patient is eligible to continue to receive non-pharmacological standard management (dependent on their sampled time to treatment discontinuation).	Clinical expert consultation suggested that standard management would continue beyond cessation of adjunctive pharmacological therapy.[31]
Weight regain	
Weight regain begins immediately after a patient discontinues all treatment (that is, adjunctive pharmacological treatment as well as standard management).	This assumption was made in the model built by Ara et al.[14] For patients who discontinue adjunctive therapy but continue to receive non-pharmacological standard management, weight regain was assumed to only commence when standard management was discontinued. Clinical expert opinion was sought to validate this assumption.[31]
Weight is regained linearly over a 3-year period.	This assumption was made in the model built by Ara et al.[14]
The regained weight is reflective of the BMI expected as predicted by the natural history model for BMI over time (Section 5.3.4.3),	BMI was assumed to revert to the natural history model predicted BMI given the intrinsic correlation known between age and BMI (as shown by the natural history model in Section 5.3.4.3). This setting was included as a scenario analysis within the report by Ara et al., but was considered the most appropriate setting within the *de novo* model for incorporating BMI over time.[14]
Obesity-related clinical events	
Within the model, it is possible for patients to experience a primary and secondary cardiovascular event (MI or stroke), as well as developing T2DM.	This assumption was made in the model built by Ara et al.[14] It is expected that the incremental clinical impact of further cardiovascular events would be negligible, as the proportion of patients who would experience more than two cardiovascular events in clinical practice is small.
Key: MI, myocardial infarction; T2DM, Type 2 diabetes mellitus; NB32, naltrexone 32mg plus bupropion.	

5.2.3 Additional model features

In addition to the key model structure and logic information presented in Section 5.2.2.3, and the key modelling assumptions presented in Section 5.2.2.4, additional

model features were incorporated to ensure the model produces rational patient outcomes.

Within the model, patients experience treatment assessment events (designated as "treatment-related events" within Figure 25) at 12 weeks, 16 weeks, 52 weeks and 56 weeks. Although some of these time points are only directly relevant to patients receiving a given treatment (e.g. NB32 patients are assessed only at Week 16 and Week 56), all patients experience each treatment-related event unless the patient has discontinued all treatment.

This aspect of the model was incorporated to minimise the risk of biasing outcomes related to one treatment in favour or against another. For example, if a patient treated with NB32 discontinued treatment at Week 15 but was not assessed at Week 12 (and therefore had no recorded weight loss compared with baseline), the model would assume they had achieved no weight loss compared with baseline. However, the same patient treated with orlistat would be recorded as having achieved a sampled weight loss compared with baseline.

Of note, this does not mean there is a consequence for response at time points un-related to the treatment received, as assessment for response is treatment-specific. NB32 patients are assessed at Week 16 and Week 56, whereas orlistat patients are assessed at Week 12 and Week 52. Further details of assessment for response are presented in Section 5.3.3.1.

Furthermore, as both NB32 and orlistat are associated with similar assessments (i.e. the only difference is that they occur at different times), outcomes at these assessments were deemed comparable. That is, weight loss for orlistat patients at Weeks 12 and 52 was assumed to be comparable to weight loss for NB32 patients at Weeks 16 and 56.

Following this, the weight loss outcomes reported for patients at primary response assessment (12/16 weeks) and secondary response assessment (52/56 weeks) were assumed within the model to be the same. For example, weight loss for orlistat patients at Week 12 was assumed the same as weight loss for orlistat patients at Week 16; and weight loss for NB32 patients at Week 12 was assumed the same as weight loss for NB32 patients at Week 16.

The assumption of equivalent weight loss at similar assessment times relates directly to the model feature that all patients experience all treatment assessment events, as described above. The same assumption was upheld within the ITC described in Section 4.10.

Table 54 summarises further features of the *de novo* economic analysis.

Table 54: Features of the *de novo* analysis

Factor	Chosen values	Justification	Reference
Time horizon	Lifetime	The model uses a lifetime horizon to reflect all important differences in costs or outcomes between the technologies being compared	NICE (2013)[115]
Were health effects measured in QALYs; if not, what was used?	QALYs	NICE reference case	NICE (2013)[115]
Discount of 3.5% for utilities and costs	Discount of 3.5% for utilities and costs	NICE reference case	Section 5.2; NICE (2013)[115]
Perspective (NHS/PSS)	NHS/PSS	NICE reference case	NICE (2013)[115]
Key: NHS, National Health Service; PSS, Personal social services; QALY, quality-adjusted life year.			

5.2.4 Intervention technology and comparators

In line with the final scope and licensed indications, the comparators for NB32 as an adjunct to standard management are (i) orlistat as an adjunct to standard management and (ii) standard management alone.

NB32 is implemented as per its EMA Summary of Product Characteristics (SmPC) posology and method of administration, incorporating a 4-week escalation period, after which the maximum recommended daily dose of 32mg naltrexone hydrochloride and 360mg bupropion hydrochloride is assumed.[29] Orlistat is similarly implemented as per its EMA SmPC posology and method of administration, a 360mg daily dose.[10]

Discontinuation rules for both NB32 and orlistat are implemented in the model, as per their license terms.[10, 29] Patients who fail to meet the response criterion of ≥5% weight loss from baseline after 12 weeks after full treatment initiation (12 weeks after initiation of post-escalation period treatment for NB32 patients; 16 weeks after treatment initiation) discontinue adjunct pharmacological therapy.

Following advice from Professor John Wilding, discontinuation rules would in practice apply 12-months after full dose initiation.[31] Patients who fail to meet the response criterion of ≥5% weight loss from baseline after 52 weeks after full treatment initiation (56 weeks after initiation of initial treatment-escalation for NB32 patients), discontinue adjunct pharmacological therapy in the analysis.

Standard management as implemented in the analysis is specified to reflect the non-pharmaceutical dietary and lifestyle management treatment received in UK NHS practice, with details provided in Section 5.5. At clinical review, Professor John Wilding advised that while what comprises standard management varies by geography, the non-pharmaceutical treatment administered in the COR-I and COR-II is a good reflection of the treatment patients are likely to receive in NHS England.[31] Resource use assumptions used by Ara et al. were used to attribute appropriate NHS resource costs to COR-I and -II standard management resources, and further validate assumptions.

As described throughout Section 5.2, Kaplan–Meier (KM) analyses of treatment duration data from the COR trial programme and NB-CVOT study datasets are used to inform accurate treatment duration assumptions in the model. These analyses are described in Section 5.3.

5.3 Clinical parameters and variables

5.3.1 Baseline patient characteristics

Baseline patient characteristics were derived from a range of sources to best represent patients in UK clinical practice. Simulation models can be used to combine multiple sources of information to elucidate and test potential solutions.[116] The baseline characteristics included within the model are shown in Table 55.

Table 55: Baseline patient characteristics

Parameter	Mean value	Justification
Age	47.0 years	COR trial programme patient-level data
Female	79.0%	
Height	Female: 1.64 m Male: 1.78 m	
BMI	Derived from model	BMI trajectory model by Ara et al.[14] (see Section 5.3.4.3)
T2DM at baseline	33.2%	Ara et al.[14]
Insulin use for T2DM patients	33.3%	Clinical opinion[31]
Smoking status	Current: 7.0% Previous: 54.0% Never: 39.0%	Dare et al.[117]
Statin use	79.3%	NB-CVOT study[52]
History of angina	0%	Assumption – no data identified for overweight/ obese patients
Other type DM	0%	

Key: BMI, body mass index; DM, diabetes mellitus; T2DM, Type 2 diabetes mellitus.

Where possible, data were utilised from the COR trial programme, followed by the NB-CVOT study and then alternative data sources. Age, gender and height values were all derived using patient-level data from the COR trial programme.

For consistency with later model projections, BMI was derived at baseline using the BMI trajectory model by Ara et al.[14] (see Section 5.3.4.3). Use of this model ensures estimated changes in BMI over time are logical, given that following all treatment cessation, patients are assumed in the base case analysis to regain weight linearly over a 3-year period until their projected BMI at this time (see Section 5.2.2.4). Average height is used to derive average weight at baseline.

The proportion of T2DM patients at baseline was taken from Ara et al.[14] as the majority of studies in NB32 and orlistat were either conducted in non-diabetics or only diabetics. Insulin use for diabetics was assumed to be 33.3%, in line with clinical expert opinion that diabetes treatment comprises of insulin for around a third of patients.[31]

The proportion of patients who currently smoke, previously smoked and have never smoked were taken from Dare et al., who conducted a cross-sectional study of nearly 500,000 UK middle-aged adults.[117] Statin use was reported within the NB-CVOT study, and as such the estimate from this study was used within the model.

Data regarding the history of angina and the proportion of patients with other type DM were unable to be identified for overweight/ obese patients. Therefore, these baseline characteristics were excluded from the model at baseline. These factors may impact the prognosis of patients; however, the directional effect of excluding these variables is unclear.

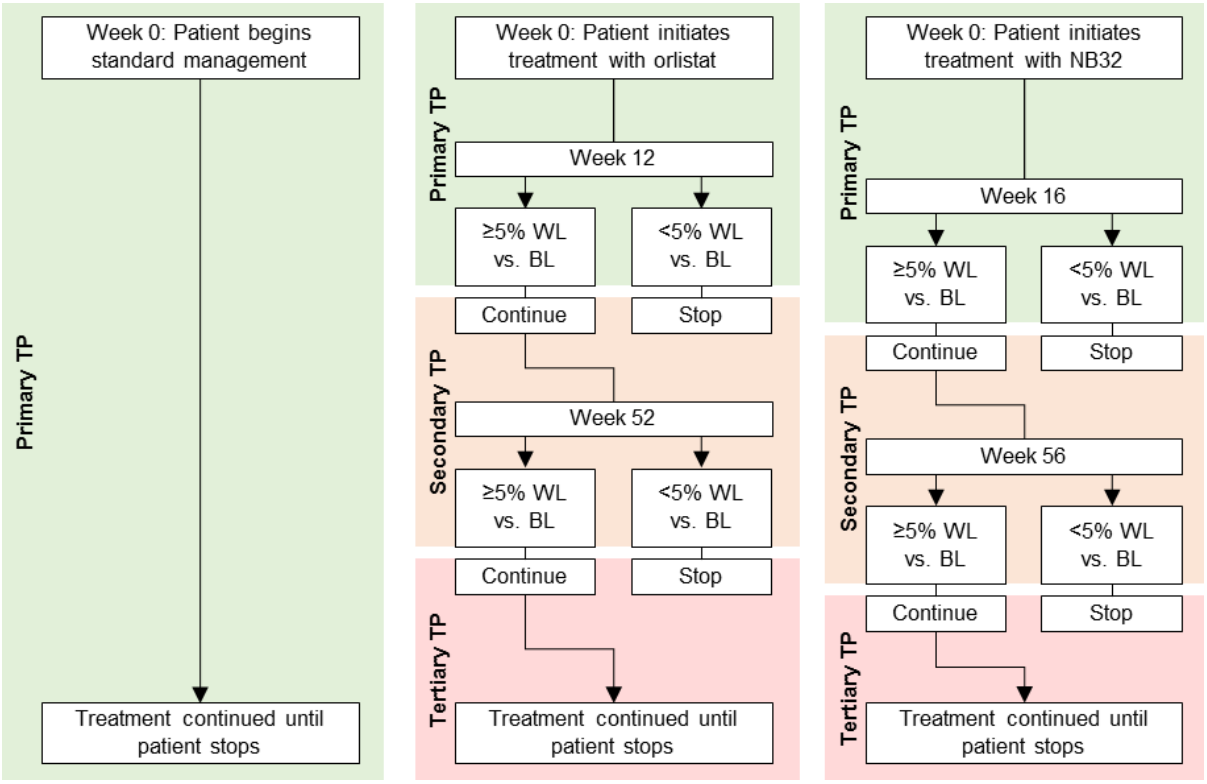
Clinical expert opinion suggested that data for the patients in the COR trial programme would be considered reflective of patients who would receive NB32 in UK clinical practice.[31]

It is recognised that the correlation between parameters in the model is not empirically considered (i.e. age, height, etc. are sampled independently of each other). However, key parameters that are directly related to others are linked appropriately (e.g. insulin use is only applied for patients sampled as T2DM, BMI is sampled in line with gender and associated height, etc.).

5.3.2 Treatment duration

The duration of treatment was applied within the *de novo* economic model in line with the expected pathway of care, as shown in Figure 26.

Figure 26: Expected pathway of care across all treatment arms



Key: BL, baseline; NB32, naltrexone 32mg plus bupropion; TP, treatment phase; WL, weight loss.

Figure 26 shows that for a patient receiving standard management, treatment is given from Week 0 until the patient stops treatment. This applies to patients receiving standard management alone, or in combination with adjunctive pharmacological therapy (i.e. NB32 or orlistat). For adjunctive treatment with NB32 or orlistat, treatment duration is considered in three phases, separated by primary and secondary assessments:

- For NB32, primary assessment is conducted at Week 16, and secondary assessment is conducted at Week 56
- For orlistat, primary assessment is conducted at Week 12, and secondary assessment is conducted at Week 52

Within the model, treatment duration data are used to determine when patients discontinue adjunctive therapy as well as standard management treatment. For adjunctive therapy, these data are used to determine when patients treated with NB32 or orlistat discontinue the adjunctive component of their treatment regimen, and revert to receiving standard management alone.

The model structure ensures that patients must cease to receive adjunctive therapy ahead of discontinuing standard management, after which they may either immediately discontinue standard management or continue to receive standard management alone. The ability for patients to continue to receive standard management following cessation of adjunctive therapy was incorporated into the model as per the clinical expert opinion of Professor Wilding.[31]

However, patients may discontinue treatment between assessment times, as KM data are used to inform the duration of adjunctive therapy and standard management treatment within the model. These data are shown for adjunctive therapy and standard management treatment in Section 5.3.2.1 and Section 5.3.2.2, respectively.

5.3.2.1 Adjunct pharmacological therapy

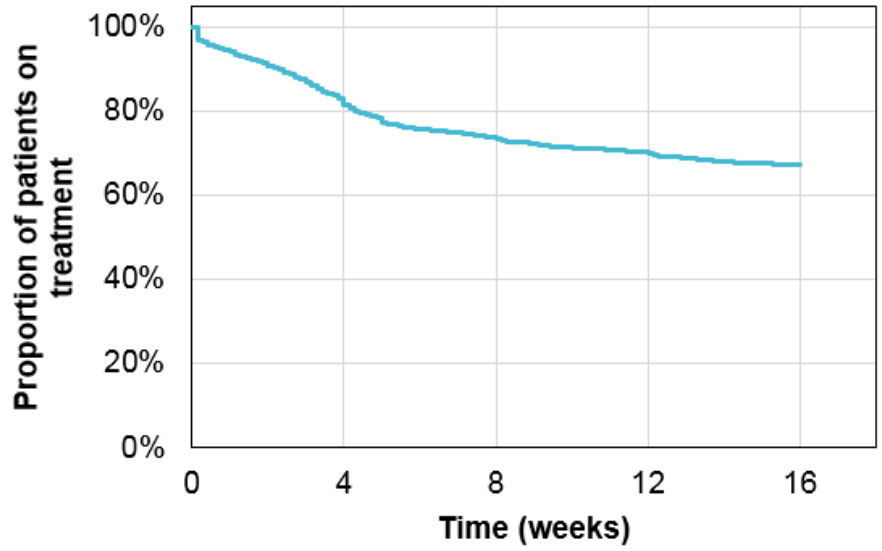
From treatment initiation to first response assessment

Patient-level data were taken from the COR trial programme and naïvely pooled. These data were subsequently analysed in SAS® to produce a KM estimate of the duration over which patients receive adjunctive therapy from treatment initiation until primary response assessment.

Figure 27 shows KM data for COR NB32 patients, up to 16 weeks. Across the COR trial programme, 67.2% of NB32 patients continued adjunct NB32 treatment until 16 weeks, and would as such have been eligible for 16-week response assessment in NHS clinical practice.

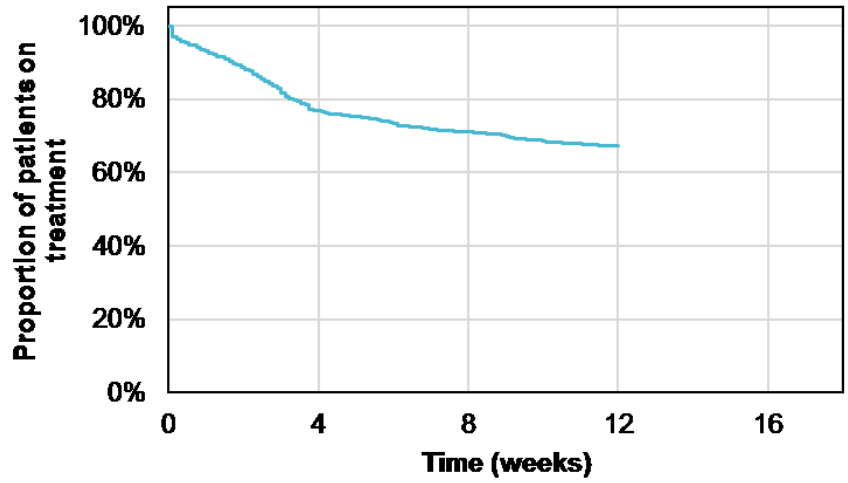
For orlistat patients, there were no comparable duration of treatment data available to inform discontinuation ahead of primary assessment reported in identified journal articles or regulatory reports. Therefore, the same KM data are used in the model to inform orlistat discontinuation assumptions up to the first response assessment. The KM data for NB32 patients were linearly scaled to fit the 12-week period to response assessment for orlistat adjunct therapy. These data are shown in Figure 28.

Figure 27: NB32 adjunct therapy discontinuation from treatment initiation to 16 weeks (pooled COR trial programme data, all NB32 patients)



At risk (COR)	2482	2027	1823	1734	1667

Figure 28: Orlistat adjunct therapy discontinuation from treatment initiation to 12 weeks (from pooled COR trial programme data, all NB32 patients)



From treatment initiation to second response assessment

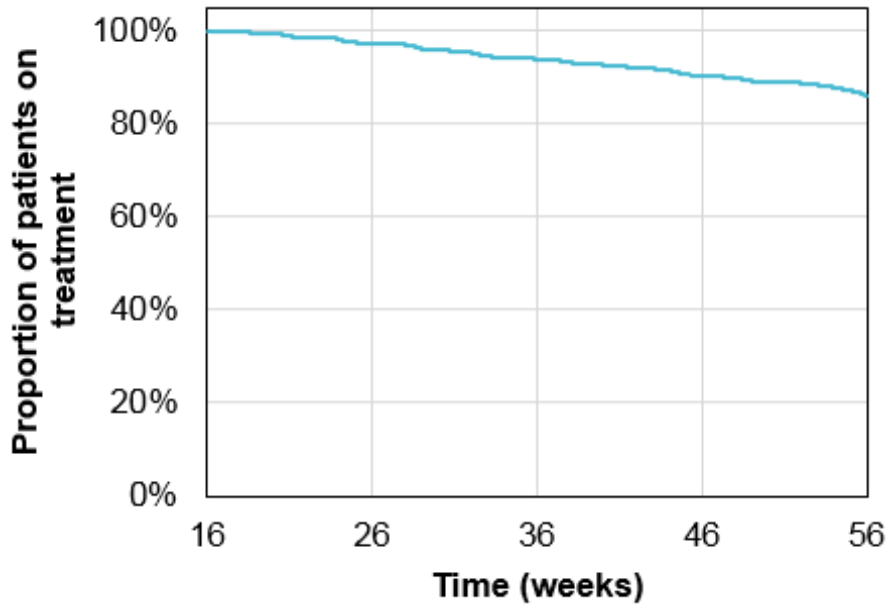
As per the analysis of the duration of adjunctive treatment in the primary phase, patient-level data were taken from the COR trial programme and naïvely pooled. Patients were included within this analysis if they achieved a weight loss of at least

5% compared with baseline at their primary assessment date. These data were subsequently analysed in SAS to produce a KM estimate of the duration over which patients receive adjunctive therapy from primary response assessment until secondary response assessment.

Figure 29 shows KM data for COR NB32 patients, from 16 weeks up to 56 weeks. Across the COR trial programme, 86.1% of responding NB32 patients (at Week 16) continued adjunct NB32 treatment until 56 weeks, and would as such have been eligible for 56-week response assessment in NHS clinical practice.

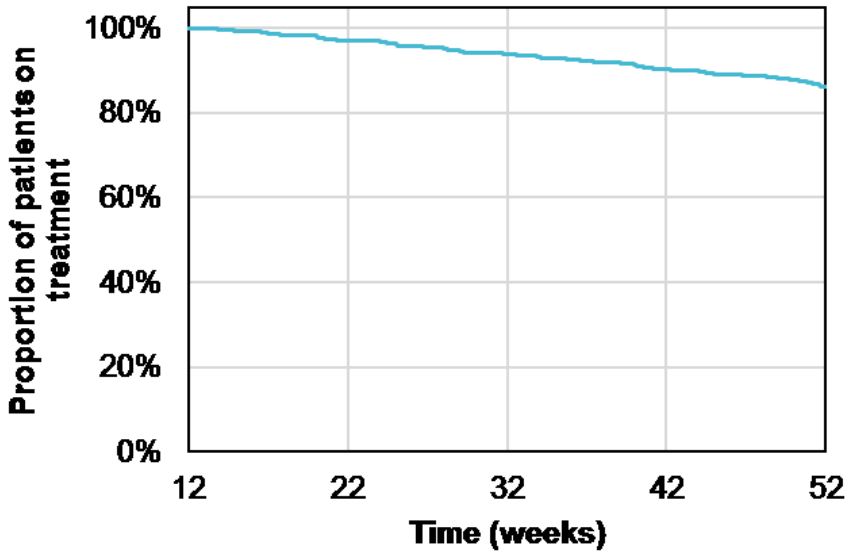
For orlistat patients, there were no comparable duration of treatment data available to inform discontinuation of responsive patients ahead of secondary assessment reported in identified journal articles or regulatory reports. Therefore, the same KM data are used in the model to inform orlistat discontinuation assumptions from the primary response assessment up to the second response assessment. The KM data for NB32 patients were transformed to match the treatment period for orlistat adjunct therapy (i.e. shifted by 4 weeks). These data are shown in Figure 30.

Figure 29: NB32 adjunct therapy discontinuation from 16 to 56 weeks (from pooled COR trial programme data; NB32 16-week responders)



At risk (COR)	1037	1007	973	936	893

Figure 30: Orlistat adjunct therapy discontinuation from 12 to 52 weeks (from pooled COR trial programme data; NB32 16-week responders)



From second response assessment onwards

Patient-level data were required to derive the duration of treatment following secondary response assessment. However, the COR trial programmes only ran up to 56 weeks.[16-19] Therefore, alternative sources were sought to implement the time on adjunctive treatment following secondary response assessment.

The Phase IIIb NB-CVOT study was a randomised, multicentre, double-blind study of 8,910 overweight or obese patients at increased cardiovascular risk treated with NB32 or placebo.[52] This study provides follow-up data for the duration of NB32 treatment up to 158 weeks after randomisation. As described in Section 5.2.1, patient-level data from the NB-CVOT study comprise the best available evidence for NB32 treatment duration assumptions beyond 1 year in clinical practice. Treatment continuation, and thus benefit, are likely to be underestimated by these data nevertheless, given the age and comorbidity profile of NB-CVOT study patients.[52]

To reflect clinical practice, NB-CVOT study patients who remained on NB32 adjunct therapy at 56 weeks were included within this analysis if they had achieved a weight loss of at least 5% from baseline. These data were analysed in SAS to produce a

KM estimate of the duration over which patients receive adjunctive therapy from secondary response assessment until treatment discontinuation.

All patients were assumed to discontinue after treatment duration data were unavailable. This assumption was deemed conservative, as after cessation of adjunctive therapy, all patients go on to receive standard management, or discontinue all treatment immediately.

Figure 31 shows KM data for included NB-CVOT study NB32 patients, from 56 weeks until treatment cessation. For orlistat patients, there were no comparable duration-of-treatment data available to inform discontinuation of responsive patients following secondary assessment reported in identified journal articles or regulatory reports. Therefore, the same KM data are used in the model to inform orlistat discontinuation assumptions from the second response assessment. The KM data for NB32 patients were transformed to match the treatment period for orlistat adjunct therapy (i.e. shifted by 4 weeks). These data are shown in Figure 32.

Figure 31: NB32 adjunct therapy discontinuation from 56 weeks (from NB-CVOT study data; NB32 56-week responders)

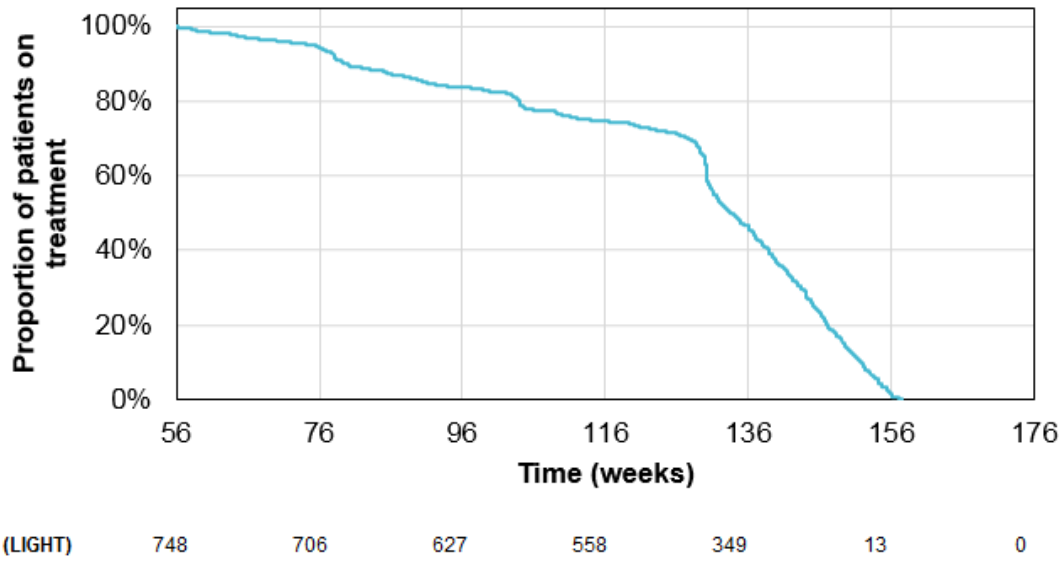
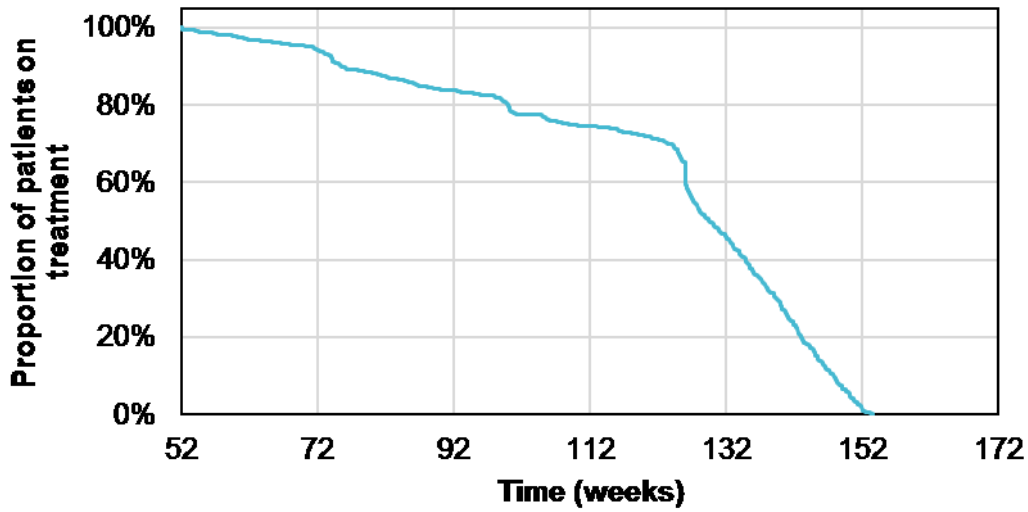


Figure 32: Orlistat adjunct therapy discontinuation from 56 weeks (from NB-CVOT study data; orlistat 52-week responders)



5.3.2.2 Standard management (non-pharmacological treatment)

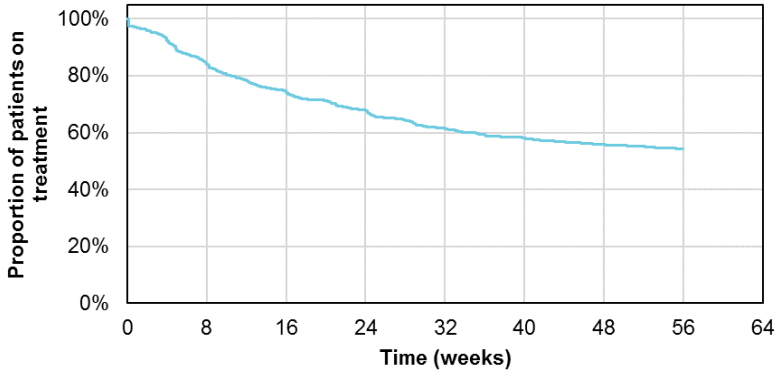
Patient-level data were required to derive the duration of treatment for patients receiving standard management. Data from the COR trial programme are available up to 56 weeks, and were therefore used to inform the model. Thereafter, data from placebo patients in the Phase IIIb NB-CVOT study were used to inform the duration of treatment from Week 56 until the end of available data in this study (approximately 158 weeks).

As patients receiving standard management alone are not subject to the same response-based treatment stopping rules as those receiving adjunctive therapy with NB32 or orlistat, all patient-level data from treatment initiation to Week 56 were utilised to inform the duration of standard management treatment. Patient-level data from NB-CVOT study placebo patients beyond 56 weeks were used to inform the duration of standard management treatment beyond the length of follow-up available in the COR trial programme.

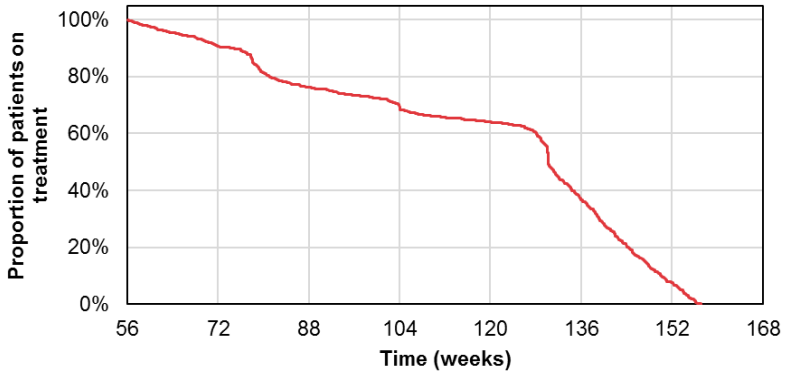
To combine these sources, KM data from the COR trial programme were used to inform treatment duration from Week 0 to Week 56. After this, KM data from the NB-CVOT study from Week 56 to Week 158 were joined to KM data from the COR trial programme by scaling the curve according to the proportion of patients who were still receiving standard management treatment at Week 56, as shown in Figure 33.

Figure 33: Derivation of duration of standard management treatment

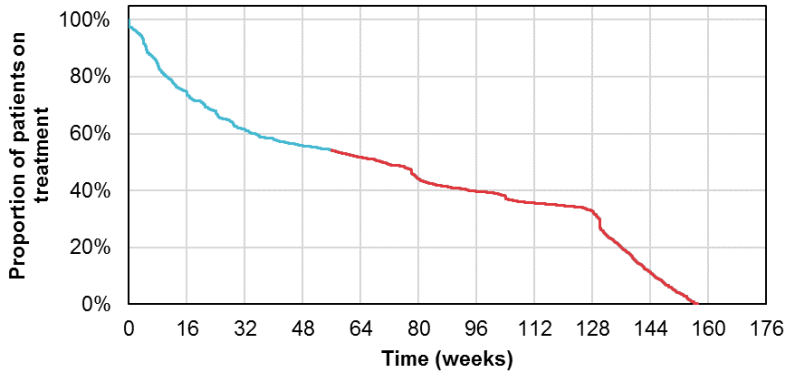
Kaplan–Meier data from the COR trial programme (Week 0 to Week 56)



Kaplan–Meier data from the NB-CVOT study (Week 56 onwards)



Combined Kaplan–Meier data from both studies (Week 0 onwards)



5.3.3 Treatment effectiveness

To implement the treatment effectiveness of NB32 and orlistat, weight loss measurements were required at key points throughout the model. These key points were defined as:

- At primary assessment (12/16 weeks)
- At secondary assessment (52/56 weeks)

Two metrics were applied in the model to account for treatment effectiveness relating to weight loss. These were:

- The proportion of patients with ≥5% decrease in body weight
- Mean change in body weight

These outcome measurements were derived using patient-level data from the COR trial programme. The implementation of these outcome measures is illustrated in Section 5.3.3.1 and Section 5.3.3.2.

5.3.3.1 Proportion of patients with ≥5% decrease in body weight

<u>At primary response assessment</u>

To account for the proportion of NB32 patients who achieved a weight loss of at least 5% compared with baseline at primary response assessment, patient-level data from the COR trial programme were used. Within the study, n=1,038 patients achieved a response at Week 16 and hence were eligible to continue treatment. Based on an estimated 67.2% of NB32 patients who continued to receive adjunct therapy until 16 weeks, and a total of N=2,043 patients randomised to receive NB32 at baseline, the proportion of patients who achieved a response at 16 weeks (given they were still receiving NB32 treatment at this time) was derived using Equation 3.

Equation 3: Proportion of NB32 patients who respond at Week 16

$$Proportion_{respond} = \frac{n}{N * 67.2\%}$$

Key: $Proportion_{respond}$, proportion of NB32 patients who respond at Week 16; n=number of 16-week responders; N, number of patients randomised to receive NB32; NB32, naltrexone 32mg plus bupropion.

Equation 3 yielded an estimated 75.7% of NB32 patients who responded at Week 16. The standard error of this sample proportion was derived using Equation 4.

Equation 4: Standard error of the sample proportion

$$SE(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$$

(where $p = \hat{p}$ if p is unknown)

Key: SE, standard error; p, known proportion; $\hat{p}$, sample proportion; n, number of responders

Equation 4 yielded an estimated standard error of 0.012.

For patients treated with orlistat, the equivalent proportion of patients who responded at primary response assessment was derived via the *de novo* ITC, reported in Section 4.10. As the ITC was conducted for T2DM patients and non-T2DM patients separately, separate results are applied from the ITC dependent on a sampled patients' baseline T2DM status.

The ITC did not produce relative effectiveness estimates for proportion of responders at primary assessment. Out of necessity, the ITC results for percentage of responders at 1 year is assumed to be generalisable to primary response assessment. To estimate the proportion of responders at 12 weeks, the ITC odds ratios for response between NB32 and orlistat were applied to the 16-week NB32 response estimate (75.7%).

This approach yields an estimated 73.0% of orlistat patients who responded at Week 12.

For patients treated with standard management alone, response-based discontinuation at Week 12 or 16 is not considered in the model, as no patients treated with standard management alone would discontinue treatment due to a lack of response in practice.

In summary, the figures presented in Table 56 were derived from the proportion of patients who achieved a weight loss of at least 5% compared with baseline at primary response assessment.

Table 56: Proportion of responsive patients at primary response assessment

Treatment		Proportion (SE)	Source
NB32		75.7% (0.012)	COR trial programme data
ORL	All patients	73.0%	ITC[a]
	T2DM	77.9%	ITC
	Non-T2DM	70.5%	ITC
SM		NA	NA

Key: ITC, indirect treatment comparison; NA, not applicable; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SE, standard error; SM, standard management; T2DM, Type 2 diabetes mellitus.
Notes: [a], The derived proportion shown here is an estimate based upon the proportion of T2DM patients at baseline.

At secondary response assessment

Mean change in body weight estimates determines the proportion of responders and non-responders at secondary response, as described in Section 5.3.3.2.

5.3.3.2 Mean change in body weight

At first response assessment

Following the primary treatment phase, weight loss for patients treated with NB32 was estimated separately for 16-week responders and 16-week non-responders. As described in Section 5.3.3.1, the separation of these estimates allows accurate estimation of the proportions of patients who continue through each treatment phase.

Patient-level data from the COR study programme were used to derive average weight loss for NB32 patients at Week 16 (i.e. primary response assessment). For Week 16 responders, average weight loss for NB32 patients was calculated as 9.4%, and for Week 16 non-responders, average weight loss for NB32 patients was calculated as 1.9%.

Weight loss at primary response assessment for responders is capped within model calculations at a minimum of 5% (due to the definition of a responder as exhibiting weight loss of ≥5% since baseline), and weight loss at primary response assessment for non-responders is capped within model calculations at a maximum of 4.99% (due to the definition of a non-responder as exhibiting weight loss of <5% since baseline).

For patients treated with orlistat, the equivalent weight losses were derived via the *de novo* ITC. As for the proportion of responders at primary assessment, the outcome of mean difference in weight loss after 1 year was assumed to be generalisable to apply at primary response assessment.

Relative effectiveness of treatments in terms of mean weight loss is not stratified by response in the analysis. This assumption was necessary as the relative effectiveness of NB32 and other treatments could not be stratified by response status in the ITC, as reported in Section 4.10. For example, the outcome of mean difference in weight loss for NB32 and orlistat patients for primary response assessment responders cannot be derived from available data.

Furthermore, the weight loss for NB32 patients at Week 16 was assumed comparable to the weight loss for orlistat patients at Week 12, given the lack of a 4-week titration period for patients treated with orlistat (and hence for NB32 patients, weight loss at Week 16 was assumed the same as weight loss at Week 12).

For Week 12 responders, average weight loss for orlistat patients was calculated as 8.6%, and for Week 12 non-responders, average weight loss for orlistat patients was calculated as 1.1%. Weight loss at Week 12 was assumed the same as weight loss at Week 16 for orlistat patients.

For patients treated with standard management alone, weight losses at Weeks 12 and 16 were derived using available COR trial programme patient-level data. These data were not separated by response as standard management patients would not be assessed for response.

For standard management patients, average weight loss at Week 12 was calculated as 2.3%, and average weight loss at Week 16 was calculated as 2.7%.

In summary, the figures presented in Table 57 were derived from average weight loss compared with baseline at primary response assessment.

Table 57: Average weight loss at primary response assessment

Treatment	Outcome	Value	Source
NB32	Primary Week 16 assessment: Responders	9.4%	COR trial programme data
	Primary Week 16 assessment: Non-responders	1.9%	
ORL	Primary Week 12 assessment: Responders (all patients)	8.6%[a]	ITC[a]

Treatment	Outcome	Value	Source
	Primary Week 12 assessment: Responders (T2DM patients)	9.2%	ITC
	Primary week 12 assessment: Responders (non-T2DM patients)	8.3%	
	Primary Week 12 assessment: Non-responders (all patients)	1.1%[a]	ITC[a]
	Primary Week 12 assessment: Non-responders (T2DM patients)	1.7%	ITC
	Primary Week 12 assessment: Non-responders (non-T2DM patients)	0.8%	
SM	Week 12: All patients	2.3%	COR trial programme data
	Week 16: All patients	2.7%	

Key: ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SE, standard error; SM, standard management; T2DM, Type 2 diabetes mellitus.
Notes: [a], The derived proportion shown here is an estimate based upon the proportion of T2DM patients at baseline.

<u>At second response assessment</u>

Following the secondary treatment phase, weight loss for patients treated with NB32 was derived at 56 weeks.

Patient-level data from the COR study programme were used to derive average weight loss for NB32 patients at Week 56 (i.e. secondary response assessment). At Week 56, average weight loss for NB32 patients was calculated as 11.7% for those who responded at Week 16, and 8.8% for all patients (regardless of response at Week 16). The figure of 11.7% is used as a baseline to make comparisons to orlistat (given the existence of the response-based treatment stopping rules for both NB32 and orlistat), whereas the figure of 8.8% is used as a baseline to make comparisons to standard management alone (given the lack of a response-based treatment stopping rule for standard management alone).

For patients treated with orlistat, the equivalent weight loss was derived via the *de novo* ITC described in Section 4.10. The weight loss for NB32 patients at Week 56 was assumed comparable to the weight loss for orlistat patients at Week 52, given the lack of a 4-week titration period for patients treated with orlistat (and hence for NB32 patients, weight loss at Week 56 was assumed the same as weight loss at Week 52). This assumption is described in further detail in Section 5.2.3.

At secondary response assessment, average weight loss for orlistat patients was calculated as 10.9%, based on the value of 11.7% for patients treated with NB32. Weight loss at Week 56 was assumed the same as weight loss at Week 52 for orlistat patients (as per the assumption held for NB32 patients, described in further detail in Section 5.2.3).

For patients treated with standard management alone, weight loss at Week 52 and 56 were also derived using the ITC. However, as patients treated with standard management alone are not subject to the same response-based treatment stopping rules as those patients treated with NB32 or orlistat, the base estimate from which the ITC was applied was taken to be the estimated weight loss for NB32 patients at Week 56 regardless of response at Week 16 (i.e. the value of 8.8%). Therefore, for standard management patients, average weight loss at Week 52 and Week 56 was calculated to be 4.5%.

In summary, the figures presented in Table 58 were derived from average weight loss compared with baseline at secondary response assessment.

Table 58: Average weight loss at secondary response assessment

Treatment	Outcome	Value	Source
NB32	Secondary Week 56 assessment	11.7%	COR trial programme data
ORL	Secondary Week 52 assessment (all patients)	10.9%	ITC[a]
	Secondary Week 52 assessment (T2DM patients)	11.5%	ITC
	Secondary Week 52 assessment (non-T2DM patients)	10.6%	
SM	Week 52/56 (all patients)	4.5%	ITC[a]
	Week 52/56 (T2DM patients)	5.6%	ITC
	Week 52/56 (non-T2DM patients)	3.9%	

Key: ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SE, standard error; SM, standard management; T2DM, Type 2 diabetes mellitus.
Notes: [a], The derived proportion shown here is an estimate based upon the proportion of T2DM patients at baseline.

Patient-specific weight loss at secondary response assessment in the economic analysis, based on patient characteristics (conditions), first-order uncertainty and the data in Table 58, determines whether the patient continues or discontinues adjunct pharmacotherapy therapy at secondary assessment. As described in Section 5.2.4,

patients who fail to meet the response criterion of ≥5% weight loss from baseline after 52 weeks after full treatment initiation (56 weeks after initiation of initial treatment-escalation for NB32 patients), discontinue adjunct pharmacological therapy in the analysis, based on Professor John Wilding's insight into clinical practice.[31]

5.3.4 Epidemiological models of natural history

In their 2012 review, Ara et al. summarised that natural history models of how changes in BMI affect the risk of major clinical events and how BMI levels change with age are needed to appropriately model the cost effectiveness of weight-loss strategies for overweight and obese patients.[14]

Ara et al. identified key limitations in previous explorations of the relationship between weight and the development of cardiovascular disease (CVD), diabetes or mortality, for the purposes of their economic analysis.[14] Previous studies had comprised of[14]:

- Cross-sectional studies only able to identify correlation
- Studies that categorised BMI, and thus unable to capture changes within categories
- Studies conducted primarily outside the UK

The analyses that Ara et al. conducted in light of these limitations, using large-scale GPRD data, were pivotal in allowing Ara et al. to use informed estimates of risk at specific levels of BMI and age, and capture the dynamic relationship between BMI and age, while controlling for confounding factors across analyses. The natural history models reported by Ara et al. are used to similar effect in the *de novo* model.

Chapter 4 of Ara et al.'s 220-page *Health Technology Appraisal* report documents the data, methods and results from their epidemiological data analyses. This report is publicly available without restrictions; as such, reporting here is succinct and refers the reader to data presented in the Ara et al. report where sensible.[14]

5.3.4.1 Ara et al. data preparation

Ara et al. accessed the GPRD in January 2011 and initially drew 100,000 individuals from the pool of GPRD patients who had three or more BMI readings of over

27kg/m^2.[14] Patient data prior to 1980 were removed, as were observations with missing dates, BMI readings during or within 6 months of a pregnancy, and BMI readings outside the range 25–60kg/m^2.[14]

Occurrence of all-cause mortality (ACM), myocardial infarction (MI), stroke and T2DM onset was identified for each individual, to allow analysis of TTE for each outcome. As complete patient data were not available, Ara et al. created separate patient cohorts for each outcome. Each cohort except the T2DM onset cohort was then subdivided into diabetic and non-diabetic cohorts, creating seven TTE cohorts in total. Each cohort consisted only of patients who were either diabetic or non-diabetic for their entire follow-up period, to reduce 'carry-over' effects from comorbidities occurring when, for example, a patient was non-diabetic but then became diabetic.[14] Ara et al. justified this as also negating the issue of a reliable diagnosis of diabetes. A patient may be diagnosed as diabetic; however, there may be a substantial lag before their GPRD record reflects this.[14]

Ara et al. documented their selection of available covariates, and report summary statistics for these in Tables 12–18 of their report.[14] Included covariates comprise variables to capture BMI; baseline age; sex; whether aspirin, statins, or blood-pressure-lowering treatment were being used; and smoking status.[14] Diabetic cohorts also included a covariate dummy for insulin use. Ara et al. used only baseline BMI in TTE analysis, justifying this based on previous large-scale studies.[14]

Table 11 and Figures 4 and 5 of the Ara et al. report summarise patient numbers and follow-up length for the seven GPRD TTE cohorts.[14] Less than 10% of patients in the diabetic ACM cohort had follow-up beyond 15 years; Ara et al. state it is therefore unwise to apply TTE results beyond this range.[14] Following Ara et al., ACM in the *de novo* model is informed by general population data described in Section 5.3.4.4. TTE assumptions beyond 15 years for MI, stroke and T2DM onset are described in Section 5.3.4.2.

5.3.4.2 BMI time-to-event analysis

Ara et al. fitted Weibull models to estimate TTE for each of the seven analysis cohorts.[14] The scale parameter of the Weibull hazard function was allowed to depend on all prepared covariates, irrespective of statistical significance, and higher-order polynomial terms of BMI and age, based on significance at the 5% level.[14] The

shape parameter of the Weibull hazard function was allowed to depend on a subset of prepared covariates, based on significance at the 5% level.[14]

Ara et al. present Weibull model results as regression coefficients and 95% CIs on the log scale, in Tables 19–22 of their report.[14] Further tables and figures are used to further illustrate results, and in particular the importance of diabetes.[14] This level of dissemination allowed us to apply deterministic TTE results accurately in the *de novo* model, but regression variance-covariance matrices were required to allow us to incorporate uncertainty around Ara et al. TTE parameter estimates into sensitivity analyses in Section 5.8.

The contact author for the Ara et al. report was contacted via email during model development, and communications were helpful and appreciated. An email request for TTE variance-covariance matrices (following other previous communication) was not replied to, although we appreciate staff leave and movement of key staff since Ara et al. publication may have been a factor.

Ara et al. investigated structural uncertainty around the assumptions implicit in their selected model structure for TTE analyses by testing an alternative model structure with flexible baseline hazard function and restricted cubic splines to model continuous terms.[14] The key findings were that the Weibull assumptions were supported to the extent that the added complexity of the more flexible structure was not warranted.[14]

Although data supporting TTE estimates beyond 15 years are few, it is not clear from Ara et al. that alternative assumptions were used beyond 15 years for MI, stroke and death.[14] In lieu of alternative data, and consistent with assumptions used in many oncology NICE TAs, Weibull TTE estimates are applied over the *de novo* model time horizon for obesity-related non-fatal events.

5.3.4.3 BMI trajectory analysis

To investigate how BMI changes with time, Ara et al. conducted multilevel modelling of the repeated measures of BMI, with age as the timescale.[14]

Ara et al. modelled BMI trajectories using the diabetic and non-diabetic ACM GPRD cohorts described in Section 5.3.4.1.[14] The cohorts were different to the TTE cohorts in that repeated measures (i.e. BMI) were not restricted to the baseline measure; all

patients who had BMI below 25kg/m^2 or above 60 at any point were excluded from BMI trajectory analyses.[14] Age at each BMI recording was calculated using date of measurement and year of birth, assuming all patients were born on 1 July in the absence of more detailed birth information.[14]

Ara et al. used exploratory trajectory plots from random patients to inform model specification, before applying multilevel models. Ara et al. investigated the need for random intercepts and slopes, and the correlation between them, through likelihood ratio tests, and the models were restricted to allow only a linear trajectory.[14] The model was adjusted for sex and the interaction between age and sex, based on statistical significance at the 5% level; age was centred at 45 years.[14]

Consistent with their reporting for TTE analysis, Ara et al. report multilevel regression parameter estimates and 95% CIs in Table 24 of their report, facilitating use in the *de novo* model. As for TTE analyses, in the absence of regression variance-covariance matrices, it was not possible to correctly incorporate uncertainty around BMI trajectory model estimates into sensitivity analyses in Section 5.8.

Example trajectory plots presented alongside model results in Chapter 4 of Ara et al. illustrate the large variation in BMI trajectories across individuals.

5.3.4.4 Mortality beyond fifteen years

As described in Section 5.3.3.1, less than 10% of patients in the diabetic ACM GPRD cohort had follow-up beyond 15 years; following Ara et al., general population mortality data are used to inform probability of death after 15 years.[14]

The latest available Office for National Statistics (ONS) interim life table data are used to incorporate age- and gender-specific UK mortality probability data from 15 years after model entry to the model maximum age of 100 years.[118]

5.4 Measurement and valuation of health effects

For overweight patients and patients with obesity, quality of life can be improved substantially by effective and sustained weight reduction. A wealth of evidence, reviewed here, demonstrates that quality of life is improved during effective treatment, through weight reduction and the associated improvements in obesity comorbidities and symptoms.

The long-term benefits of effective weight reduction are manifold, indirect and inherently difficult to capture in economic analyses. Analyses of Health Survey for England data have estimated the joint effect of weight and weight-related conditions upon EQ5D utility, and these are harnessed for use in the economic evaluation. Yet even these analyses underestimate the HRQL effect of weight gain, as the relationship between utility and many known weight- and treatment-related conditions are not captured.

5.4.1 HRQL data from clinical trials

HRQL was assessed in patients in the four pivotal, multicentre, randomised, double-blind, placebo-controlled, Phase III studies in the COR trial programme, using the IWQOL-Lite questionnaire. Aside from this, patients completed the SF-36 questionnaire in COR-II only, at baseline, Week 28 and Week 58.

The IWQOL-Lite assesses the impact of weight on quality of life in five domains: physical function; self-esteem; sexual life; public distress; and work.[119, 120] Section 4.7 summarises IWQOL-Lite outcomes from the pivotal Phase III studies. Relative improvements in each of the quality of life domains were observed in NB32 patients relative to placebo patients, and overall IWQOL-Lite score improvements were observed for NB32 versus placebo patients across the three studies with diabetes exclusion criteria.

While there are many advantages of IWQOL-Lite being a weight-change specific instrument that is designed to focus on the domains, characteristics, and complaints most relevant for weight-loss-targeting patients, for appraisal of the value of NB32 for NHS England, a generic measure of HRQL is required, for fair resource allocation. Specifically, the NICE reference case documents a stated preference for patient-reported EQ-5D data.[115] Although SF-36 data were collected in COR-II, the frequency of completion and limited follow-up of the COR trials limit the usefulness of these data for the purpose of the economic analysis. As such, the available patient EQ-5D data from the literature are pivotal to address the decision problem.

5.4.2 HRQL studies

The systematic search for HRQL studies targeted MEDLINE, MEDLINE In-Process, Embase, EconLit, NHS EED and the CRD-HTA database. Eligibility criteria for the review are described in Table 59.

Further details of the search strategy are presented in Appendix 16, alongside details of the study filtering and data extraction processes.

Table 59: Eligibility criteria for the HRQL evidence search

Criteria	Inclusion	Exclusion
Population	• Adults who are obese (BMI ≥30kg/m^2), or overweight (BMI ≥25kg/m^2, adopting the most inclusive criterion from summary of product characteristics and care guidelines: that used in NICE Clinical Guideline 189) with one or more comorbidities (T2DM, dyslipidaemia and/or controlled hypertension)	• Healthy volunteers • Children (age <18 years) • Diseases other than that specified in inclusion criteria
Intervention/ comparator	• No specific inclusion criteria • Studies reporting utility values for non-treated patients will also be included to assess the burden of illness	• Studies will not be excluded based on intervention/ comparator
Outcomes	• Utility values	
Study types	• Economic evaluations reporting utility values • RCTs and observational studies reporting utility data • Studies must present sufficient detail regarding the methodology used • Studies must provide extractable results	• Non-systematic reviews[a], letters, comment or editorials • Studies not reporting adequate methodology or extractable data
Language	• Studies published in English will be included • Studies published in non-English languages will be included and flagged[b]	

Key: BMI, body mass index; NICE, National Institute for Health and Care Excellence; RCT, randomised controlled trials; T2DM, Type 2 diabetes mellitus.
Note: [a], Systematic reviews will be included and flagged for bibliography searches; [b], Studies published in languages other than English will be explored only if sufficient evidence is not identified from English studies.

The PRISMA diagram in Appendix 16 presents the flow diagram of studies identified for the HRQL review. Database searching identified 2,448 citations, with three

additional citations identified through bibliographic searching and six abstracts identified from conference proceedings.

Following screening and eligibility assessment, 49 publications were identified from which a total of 39 studies were included in the review.[14, 121-158] A tabular summary of the characteristics of each included study is provided in Appendix 16.

Twenty-five of the 39 included studies reported EQ-5D data. These studies varied in terms of geographical location, methodology, patient characteristics, making synthesis and comparison difficult. A proportion of these studies aimed to identify a relationship between BMI and EQ-5D utility, and there is a wealth of evidence that BMI is negatively correlated with utility in overweight and obese patients.[132, 137, 139, 140, 146, 149, 153, 157] The review also brings to light evidence on the relationship between utility and weight-related comorbidities in overweight and obese patients. Different studies have reported data from obese and overweight patients with different weight-related comorbidities, including: diabetes, hypertension and hyperlipidaemia[150]; joint and spinal complaints[126, 128, 134]; coronary heart disease[153]; and multiple sclerosis.[152]

The inability to explain the interrelated importance of both weight and weight-related comorbidities for patient utility limits the usefulness of most included studies for this appraisal. Ara et al. (2012)[14], described through Sections 5.1 and 5.2, informed their model utility assumptions using analysis of a large sample of data from the HSE database. The relationship between EQ-5D utility and BMI was analysed, controlling for heart disease, stroke and diabetes status variables, as well as age and gender. Results from these analyses were directly relevant for the health states captured in the Ara et al. model, and as such, suited for the *de novo* economic analysis developed for this submission.

Ara et al. reported fitting an adjusted censored mixture model (ACMM), a model structure suited to non-normally distributed and censored data[159], to historic HSE patient EQ-5D data.[14] However, disseminated details of methods and results comprised two paragraphs in Chapter 4 of Ara et al., and two tables in Chapter 5 of Ara et al. The first of these tables (Table 34, Ara et al.) shows results from the ACMM regression, and the second (Table 35, Ara et al.) shows a selection actual versus predicted scores for different plausible patient characteristic combinations. The regression results from Table 34 of Ara et al. were incorporated into the *de novo*

model. However, the implied utility values for patients in the *de novo* model were not consistent with Table 35 of Ara et al., suggesting incorrect implementation. In lieu of further information or the original model, this could not be resolved.

The PHE weight management economic assessment tool, identified in the search for previous economic analyses and described in Section 5.1, also uses results from regression analysis, of individual-level EQ-5D data drawn from HSE from 2011 to 2013, to inform HRQL assumptions.[112] Results from Tobit and ordinary least squares (OLS) regression analyses of these data are shown in Table 60. Coefficient estimates were obtained from the tool itself; further details including variance-covariance estimates were helpfully provided by the PHE tool author, Dr Vicky Copley, in response to an email request.

The models specified by the PHE team includes explanatory variables for BMI, age, gender, and the obesity-related conditions in the *de novo* model, and are therefore well suited to inform utility assumptions in the model. The results in Table 60 suggest that BMI has an independent and inverse relationship with BMI, consistent with evidence from other studies[132, 137, 139, 140, 146, 149, 153, 157], and that stroke, MI and T2DM are important for HRQL, as expected.

The model includes a covariate for cancer; colorectal and breast cancers are captured in the PHE tool.[112] While cancers are not considered to be weight-related in the *de novo* model structure, this poses no problem for implementation; simulated individuals are assumed to be cancer-free in utility calculations. The PHE tool uses World Obesity Federation relative risk estimates for colorectal and breast cancer for people with BMI of 22 or above[160] to inform the link between weight and cancer risk.[112] While it is difficult to estimate the relationships between BMI and related diseases, the inclusion of cancer as a weight-related condition in the PHE tool and absence of it in the *de novo* model (following Ara et al.) further illustrates how the *de novo* analysis inherently underestimates the benefits of weight reduction.

Table 60: Public Health England weight management economic assessment tool v2 HSE EQ-5D data analysis

Covariate	Coeff.	Variance-covariance matrix									
		BMI	BMI^2	BMI^3	Age	Female	Stroke	MI	Cancer	T2DM	Const.
Tobit Model Estimates[a]											
BMI	0.05911	0.00008									
BMI^2	-0.00175	0.00000	0.00000								
BMI^3	0.00001	0.00000	0.00000	0.00000							
Age	-0.00440	0.00000	0.00000	0.00000	0.00000						
Female	-0.04054	0.00000	0.00000	0.00000	0.00000	0.00002					
Stroke	-0.18280	0.00001	0.00000	0.00000	0.00000	0.00001	0.00059				
MI	-0.16122	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00048			
Cancer	-0.16403	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00003	-0.00003	0.00028		
T2DM	-0.11093	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00002	0.00001	0.00001	0.00012	
Constant	0.67263	-0.00084	0.00002	0.00000	0.00000	-0.00008	-0.00010	0.00006	0.00002	-0.00001	0.00940
Ordinary Least Squares Regression Estimates											
BMI	0.03293	0.00003									
BMI^2	-0.00094	0.00000	0.00000								
BMI^3	0.00001	0.00000	0.00000	0.00000							
Age	-0.00219	0.00000	0.00000	0.00000	0.00000						
Female	-0.02258	0.00000	0.00000	0.00000	0.00000	0.00001					
Stroke	-0.12652	0.00000	0.00000	0.00000	0.00000	0.00000	0.00044				
MI	-0.11931	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00035			
Cancer	-0.10944	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00002	-0.00001	0.00017		
T2DM	-0.07800	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00001	0.00001	0.00000	0.00007	
Constant	0.65792	-0.00028	0.00001	0.00000	0.00000	-0.00002	-0.00004	0.00001	-0.00002	-0.00002	0.00311

Key: BMI, body mass index; EQ-5D, EuroQol 5-dimensions; HSE, Health Survey for England; MI, myocardial infarction; OLS, ordinary least squares; T2DM, Type II diabetes mellitus.

Notes: [a], Censoring limits were -0.594 and 1; sigma 0.33898 (standard error 0.00365) (both to 5 decimal places).

5.4.3 Adverse reactions

As described in Section 4.12, the safety profile of NB32 is consistent with its individual drug components, and different to the tolerability profile of orlistat. At clinical review, Professor Wilding expressed a belief that NB32 patients have a HRQL benefit over orlistat patients as a result of AE differences.[31] While the side effects associated with NB32 are similar to those associated with many common drugs, the lower digestive tract AEs associated with orlistat can be particularly unpleasant for patients.[31] In addition, as described in Section 4.12, while no NB32-related deaths were observed across the COR trial programme and the NB-CVOT and IGNITE studies, an orlistat mortality risk from increased liver reaction risk cannot be ruled out based on clinical study data.

The retrospective regulatory stopping rules for NB32 and orlistat, while limiting unnecessary drug exposure and therefore limiting adverse reactions, make estimating comparative AE profiles beyond 16 weeks very difficult without access to patient-level data from key trials for both NB32 and orlistat. Details of orlistat AEs from the clinical trial literature and publicly available regulatory documents are not sufficient to make appropriate trial-data comparisons between NB32 and orlistat adjunct therapies.

Reporting of AEs was varied across orlistat studies. The pivotal trial publication of the largest orlistat study identified (XENDOS) by Torgerson et al. did not report specific AEs.[83] Overall, the reporting of orlistat AE severity was scant, and almost non-existent for AE duration.

Aside from these clinical data problems, the HRQL implications of the orlistat and NB32 AEs for obese and overweight patients are poorly understood, and in some cases overlap and interact with obesity comorbidities. For example, within the COR trial programme, anxiety and depression AEs were recorded, but it is unclear how many of these incidences are treatment-dependent or condition-related. The *de novo* model uses COR trial programme AE incidence data and assumptions to account for AE costs, as described in Section 5.5.4.

For AE HRQL effects, the model assumes no on-treatment differences across treatment arms other than those indirectly implied by changes in BMI and obesity-related disease, from the data in Table 60. This simplifying assumption is sensible

given the data limitations. Furthermore, it is important to note that the assumption is conservative, for both the comparison to orlistat and the comparison to standard management. Compared with orlistat, NB32 patients are expected in practice to have superior HRQL to orlistat patients, owing to treatment effectiveness and relative AE profiles.[31] Although Section 4.12 shows the treatment-related AE profile of NB32 plus standard management to be worse than that for standard management alone, as documented in Section 5.4.1, NB32 patients reported overall IWQOL-Lite score improvements versus placebo patients across COR-I, COR-II and COR-BMOD. These data suggest the direct treatment benefits of NB32 adjunct therapy outweigh any AE HRQL effects attributable to NB32. In clinical trials and in practice, treatment-related AEs are generally resolved quickly, with only short-term effects upon HRQL.

5.4.4 HRQL data used in cost-effectiveness analysis

The HRQL data used in the base case cost-effectiveness analysis are the results from PHE Tobit regression analysis of recent HSE EQ-5D individual-level data reported in Table 60 of Section 5.4.2. Results from OLS regression analysis of these data, also shown in Table 60, are used in an alternative scenario explored in Section 5.8.3. Throughout this section, it has been illustrated how this approach, although based on the best available data for the model, is inherently conservative, in (i) assuming the weight-related clinical events with HRQL implications are restricted to cardiovascular events and T2DM onset and (ii) assuming no on-treatment utility differences across treatment arms apart from those captured via BMI and weight-related clinical conditions in Table 60.

5.5 Cost and healthcare resource use identification, measurement and valuation

5.5.1 Resource identification, measurement and valuation studies

The systematic literature search for resource identification, measurement and valuation studies was run alongside the search for published cost-effectiveness and HRQL studies, targeting MEDLINE, MEDLINE in process, Embase, Cochrane Library, NHS EED and CRD HTA. A detailed search strategy is provided in Appendix 17.

Inclusion and exclusion criteria for the review are described in Table 61. Further details of the search strategy are presented in Appendix 17, alongside details of the study filtering and data extraction processes.

Table 61: Eligibility criteria for the cost and resource use evidence search

Criteria	Inclusion	Exclusion
Population	• Adults who are obese (BMI ≥30kg/m^2), or overweight (BMI ≥25kg/m^2, adopting the most inclusive criterion from summary of product characteristics and care guidelines: that used in NICE Clinical Guideline 189) with one or more comorbidities (T2DM, dyslipidaemia and/or controlled hypertension)	• Healthy volunteers • Children (age <18 years) • Diseases other than that specified in inclusion criteria
Intervention/ comparator	• No specific inclusion criteria	• Studies will not be excluded based on intervention/comparator
Outcomes	• Cost data (direct or indirect, unit or total) • Resource use data • Cost of management of treatment related adverse events	• Studies will not be excluded based on outcomes
Study types	• Observational studies reporting cost and resource use data • Economic evaluations reporting cost and resource use data	• Non-systematic reviews[a], letters, comments and editorials • Studies reporting clinical data only will be excluded
Language	• Studies published in English will be included • Studies published in non-English languages will be included and flagged[b]	
Country	• Studies reporting cost and resource use data for relevant UK population will be included	• Non-UK studies will be excluded
Publication timeframe	• Studies published in or after 2006 (last 10 years)	• Studies published before 2006

Key: BMI, body mass index; CMs, comorbidities; T2DM, Type 2 diabetes mellitus.
Note: [a], Systematic reviews will be included and flagged for bibliography searches; [b], Studies published in languages other than English will be explored only if sufficient evidence is not identified from English studies.

The PRISMA diagram in Appendix 17 presents the flow diagram of studies identified for the cost and resource use review. In total, 1,510 citations were identified through database searching, with two additional citation identified through bibliographic searching and three abstracts identified from conference proceedings.

Following screening and eligibility assessment, 22 publications were identified from which 20 studies were included in the review.[14, 108, 111, 127, 161-176] A tabular summary of the characteristics of each included study is provided in Appendix 17.

Across studies, most non-pharmaceutical treatment costs were dietitian consultation and psychologist visit costs. Aside from these, the cost burden of weight-related diseases was a feature of the review. One study reported a high cost burden for obesity-related cancers that increases with BMI[173], another estimated that the average total cost of prescription medication increased with BMI[174], while a further study investigated the rate of hospital admissions in middle-aged women and estimated that one in eight hospital admissions can be attributed to overweightness or obesity.[164]

The level of reporting was generally poor across studies, to the extent that it was difficult to elicit resource use estimates in a form useful for this analysis. A notable exception to this was the Ara et al. study[14], which has informed other aspects of the analysis documented through Sections 5.1, 5.2, 5.3 and 5.4. Data from this study were particularly useful in informing healthcare resource use assumptions in the *de novo* analysis, as described throughout this section.

5.5.2 Intervention and comparators costs and resource use

Table 62 summarises the drug acquisition costs associated with NB32 and orlistat. The price presented within this submission for NB32 (8mg naltrexone/90mg bupropion), of £73.00 per pack of 112 tablets, is the price submitted to the Department of Health.

NB32 is associated with a 4-week titration period over which the dosage increases from one tablet per day to four tablets per day.[29] The dosage for the titration period and beyond is as follows:

- Week 0: One tablet in the morning, every day
- Week 1: One tablet in the morning and one tablet in the evening, every day

- Week 2: Two tablets in the morning and one tablet in the evening, every day
- Week 3 onwards: Two tablets in the morning and two tablets in the evening, every day

The cost of orlistat 120mg is £18.44 per pack of 84 capsules.[177] Orlistat does not have an associated titration period, the dose from Day 1 is three capsules daily. Evidence shows that branded version of orlistat (Xenical) accounted for less than 1% of the total prescription items for orlistat in 2015.[178] Therefore, costs for Xenical are not included.

There are no drug costs associated with standard management.

Table 62: Drug acquisition costs

Treatment	Pack size	Cost per pack	Cost per tablet	Source
NB32	112	£73.00	£0.65	List price submitted to the Department of Health
ORL	84	£18.44	£0.22	MIMS[177]
Key: MIMS, Monthly Index of Medical Specialities; NB32, naltrexone 32mg plus bupropion; ORL, orlistat.				

For completeness, the model includes settings for drug administration costs. However, as both NB32 and orlistat are oral medicines, it is anticipated that there are no costs associated with their administration. The model allows the user to manually input administration costs if necessary.

Table 63 presents the administration costs applied in the model.

Table 63: Administration costs

Treatment	Administration cost	Source
NB32	£0.00	No cost
ORL	£0.00	No cost
SM	£0.00	No cost
Key: NB, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.		

The medical resource use items comprising “standard management” in the model are GP visits, nurse visits and blood tests. Table 64 and Table 65 show these items, their associated costs and expected frequencies for the population considered.

Table 64: Medical resource use item costs

Resource	Cost	Source
GP visit	£44.00	PSSRU (2015) – Per patient contact lasting 11.7 minutes, including direct care staff costs, with qualification costs[179]
Nurse visit	£14.47	PSSRU (2015) – Per patient contact lasting 15.5 minutes, including qualifications[179]
Blood test	£3.01	NHS reference costs (2015) – Code DAPS05[180]
Key: DAPS, Direct Access Pathology Services; GP, general practitioner; NHS, National Health Service; PSSRU, Personal Social Services Research Unit.		

The composition of non-drug resource use was determined based on a combination of reporting in the COR studies[16-19], the publication by Ara et al.[14] and UK clinical expert opinion.[31] Non-drug resource use in the COR trial programme is detailed in Table 65.

Table 65: Non-drug resource use in COR Phase III trials

Trial	Non-drug resource use
COR-I[16]	Participant assessments were undertaken at screening and every 4 weeks. At baseline and at 12, 24, 36, and 48 weeks, participants in each group were instructed to follow a hypocaloric diet (500 kcal per day deficit based on the World Health Organisation [WHO] algorithm for calculating resting metabolic rate) and were given advice on lifestyle modification (including instructions to increase physical activity).
COR-II[17]	Study visits occurred at baseline and every 4 weeks. At baseline, 12, 24, 36, and 48 weeks, participants received instructions to follow a hypocaloric diet (500 kcal/day deficit) and increase physical activity, and behavioural modification advice.
COR-DM[19]	Participant assessments were undertaken at screening, baseline, and every 4 weeks thereafter. At baseline and Weeks 4, 16, 28, and 40, all participants were instructed by study site personnel to follow a hypocaloric diet (500 kcal deficit/day, based on the WHO algorithm for calculating resting metabolic rate).
COR-BMOD[18]	All participants in both treatment groups received an intensive program of BMOD that was delivered to groups of 10–20 persons by registered dietitians, behavioural psychologists, or exercise specialists. Group meetings lasted 90 minutes (including the weigh-in) and were held weekly for the first 16 weeks, every other week for the next 12 weeks, and monthly thereafter (yielding a total of 28 sessions).

Table 66 illustrates the non-drug (standard management) treatment assumptions in the analysis. During the clinical validation meeting, Professor Wilding verified that the non-drug treatment received alongside NB32 in the COR-I and COR-II clinical trials is a good reflection of the average diet and exercise regimens prescribed for obese and overweight patients in the UK.[31] It was added that these could be delivered by dietician, GP or Weightwatchers, dependent on postcode.[31] It is for this reason that we have included five GP visits in the first year for all treatment arms. The only difference between the two active treatments is the timing of the visits: patients receiving NB32 are assessed at Week 16 to determine treatment continuation in line with the EMA stopping rule.[29] For orlistat patients, this assessment is at Week 12.[10] Similarly, in line with clinical expectation, after receiving a full year of treatment, patients are reassessed to determine whether the initial 5% weight loss has been maintained; this occurs at 52 weeks for orlistat and 56 weeks for NB32. It has been assumed in the model that the need for continued treatment is reassessed annually by a GP.

The study by Ara et al.[14] assumed that all weight management patients had monthly visits to see a healthcare professional. In addition, patients on active adjunct treatments received blood tests at baseline and 3 months. These assumptions have been incorporated into the *de novo* model.

In addition, based on UK clinical expert consultation, all weight management patients would have an annual blood test to monitor blood glucose levels. For NB32 patients this is costed once per year at Week 56. For orlistat and standard management patients, this is costed once per year at Week 52 and would occur at the same time as the annual GP appointment. While the patient is still on treatment, it is assumed that the monthly surgery visits would continue. Therefore, visits with the practice nurse have been costed every 4 weeks from Week 60 onwards.

In line with clinical expert opinion, patients receiving standard management alone would incur approximately the same non-drug resource use costs as patients receiving adjunctive therapy alongside standard management (excluding additional blood tests for patients receiving adjunctive therapy).

Table 66: Non-drug treatment assumptions in the *de novo* model

Time (weeks)	NB32			ORL			SM		
	GP	Nurse	Blood	GP	Nurse	Blood	GP	Nurse	Blood
0	1	0	1	1	0	1	1	0	0
4	0	1	0	0	1	0	0	1	0
8	0	1	0	0	1	0	0	1	0
12	0	1	0	1	0	1	1	0	0
16	1	0	1	0	1	0	0	1	0
20	0	1	0	0	1	0	0	1	0
24	1	0	0	1	0	0	1	0	0
28	0	1	0	0	1	0	0	1	0
32	0	1	0	0	1	0	0	1	0
36	1	0	0	1	0	0	1	0	0
40	0	1	0	0	1	0	0	1	0
44	0	1	0	0	1	0	0	1	0
48	1	0	0	1	0	0	1	0	0
52	0	1	0	1	0	1	1	0	1
56	1	0	1	0	1	0	0	1	0
60+[a]	0	1	0	0	1	0	0	1	0

Key: GP, general practitioner; NB, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

Notes: [a], These frequencies apply from Week 60 every 4 weeks while patients are still receiving treatment.

Based on Table 64 and Table 66, a summary of costs was produced for a year of full treatment. Non-drug treatment costs for patients completing 1 year of treatment are £403, £403 and £397 for NB32, orlistat and standard management patients, respectively.

5.5.3 Health-state and condition-specific resource use and unit costs

Costs associated with obesity-related comorbidities were sourced from Ara et al.[14] and adapted following UK clinical expert consultation.

Costs reported by Ara et al. were inflated from 2009 levels to 2015 levels, using the Personal and Social Services Research Unit (PSSRU) Hospital and Community Health Services (HCHS) index.[179] Ara et al. incorporated costs for: cost of MI (Year

1), MI (Year 1+), stroke (Year 1), stroke (Year 1+), T2DM (Year 1), MI plus T2DM (Year 1), MI plus T2DM (Year 1+), stroke plus T2DM (Year 1), stroke plus T2DM (Year 1+), fatal stroke and fatal MI.

Ara et al. estimated health-state and condition-specific costs using the available literature.[14] The cost of MI in the first, and in subsequent years, was taken from a previous economic evaluation of early high-dose lipid lowering therapy to avoid cardiac events, which used bottom-up costing methods and considered hospitalisation, procedural, medical resource use and drug costs.[181] The cost for fatal MI was taken from a HTA evaluating the cost effectiveness of glycoprotein IIb/IIIa antagonists in non-ST elevation acute coronary syndrome.[182] The cost of stroke in the first year and in subsequent years, as well as fatal stroke, was taken from a UK study that used weighting methods taking into account the proportion of patients experiencing mild, moderate and severe strokes, in addition to discharge location.[183] Costs for T2DM with and without concomitant cardiovascular disease were taken from the literature.[184, 185]

Consultation with Professor Wilding confirmed that the NHS costs associated with MI, stroke and T2DM can be assumed to be additive.[31]

It is not clear whether Ara et al. incorporated T2DM costs after the first year of onset. To account for the cost of diabetes, a report summarised by Diabetes UK was used.[186] The report estimated monitoring and medication costs to be between £300 and £370 per patient per annum. These costs are reported for Type 1 and Type 2 diabetic patients without stratification and were used in the absence of specific Type 2 data. However, as Type 1 diabetics make up a small minority of cases, this is unlikely to be an issue.[37] Within the model, an average of these two estimates (£335) was used. As the report was written in 2012, the costs reported were inflated to 2015 levels.[179]

Ara et al. included a cost upon death, if the death was caused by MI or stroke. The figure for CVD mortality as a proportion of overall mortality (31%) was taken from WHO 2016 data.[187] Of the deaths attributable to CVD, the proportions of deaths caused by MI (43.1%), stroke (32.9%) or other causes (24.0%) were taken from WHO 2004 data.[188] From this information, mortality related to MI and stroke, as a proportion of overall mortality, was calculated as 13.4% and 10.2%, respectively.

Table 67: Medical resource use costs for comorbidities

Category	Cost	Source
MI (Year 1)	£4,210.75	Ara et al.[14] (costs inflated using PSSRU HCHS inflation indices)[179]
MI (Year 1+)	£345.91	
Stroke (Year 1)	£9,482.78	
Stroke (Year 1+)	£2,664.16	
T2DM (Year 1)	£347.57	Diabetes UK (2016)[186] (costs inflated using PSSRU HCHS inflation indices)[179]
T2DM (Year 1+)	£347.57	
Fatal stroke	£8,671.94	Ara et al.[14] (costs inflated using PSSRU HCHS inflation indices)[179]
Fatal MI	£1,390.80	
Key: T2DM, Type 2 diabetes mellitus; HCHS, Hospital and Community Health Services; MI, myocardial infarction; PSSRU, Personal Social Services Research Unit.		

Section 5.1 highlighted evidence of a positive correlation between increasing BMI and the risk of developing 17 different cancers[173, 189] and increases in prescription costs or hospital admissions for overweight and obese individuals.[164, 174] In only capturing the downstream costs of T2DM, stroke and MI, the model is inherently conservative in its ability to capture the full benefit of NB32 adjunct therapy.[31]

5.5.4 Adverse reaction unit costs and resource use

AE rates for patients on NB32 and standard management were taken from the largest of the COR Phase III trials: COR-I. Costs were included for all AEs that occurred in at least 5% of patients in either treatment arm, regardless of severity. These criteria were selected to reflect British National Formulary criteria of all very common and the majority of common AEs.[190]

The base case assumes AEs are treated solely within primary care at the cost of a single GP visit (Table 64). Outpatient visits were costed according to disease area using 2015 NHS Reference Costs.[180] The costs associated with each AE in the model are presented in Table 68.

Table 68: Outpatient adverse event costs

Adverse event	Cost	NHS reference costs (2015) Outpatient attendance service code
Anxiety	£241.52	710: Adult Mental Illness
Constipation	£135.18	301: Gastroenterology
Depression	£241.52	710: Adult Mental Illness
Diarrhoea	£135.18	301: Gastroenterology
Dizziness	£94.36	120: ENT
Dry mouth	£94.36	120: ENT
Headache	£175.76	400: Neurology
Hot flush	£132.75	502: Gynaecology
Insomnia	£241.52	710: Adult Mental Illness
Nasopharyngitis	£94.36	120: ENT
Nausea	£158.43	300: General Medicine
Sinusitis	£94.36	120: ENT
Upper respiratory tract infection	£135.18	301: Gastroenterology
Vomiting	£135.18	301: Gastroenterology
Key: ENT, ear, nose and throat; NHS, National Health Service.		

Instantaneous AE rates were calculated for NB32 and standard management considering the proportion of patients who suffered from each AE, and the average duration of treatment for patients in the COR-I trial (35.52 weeks for the NB32 group and 36.05 weeks for the standard management group). These rates were applied to patients in the model. AE rates for NB32 and standard management are presented in Table 69 and Table 70, respectively.

Table 69: NB32 instantaneous adverse event rates – COR-I Phase III trial

Adverse event	N (Total N=573)	Probability (within study)	Instantaneous rate	Cost
Anxiety	9	0.0157	0.000446	£44.00
Constipation	90	0.157	0.00481	£44.00
Depression	3	0.00524	0.000148	£44.00
Diarrhoea	26	0.0454	0.00131	£44.00
Dizziness	54	0.0942	0.00279	£44.00
Dry mouth	43	0.0750	0.00220	£44.00

Adverse event	N (Total N=573)	Probability (within study)	Instantaneous rate	Cost
Headache	79	0.138	0.00418	£44.00
Hot flush	30	0.0524	0.00151	£44.00
Insomnia	43	0.0750	0.00220	£44.00
Nasopharyngitis	29	0.0506	0.00146	£44.00
Nausea	171	0.298	0.00998	£44.00
Sinusitis	30	0.0524	0.00151	£44.00
Upper respiratory tract infection	57	0.0995	0.00295	£44.00
Vomiting	56	0.0977	0.00290	£44.00
Total adverse event cost per week:				£1.69

Table 70: Standard management instantaneous adverse event rates – COR-I Phase III trial

Adverse event	N (Total N=569)	Probability (within study)	Instantaneous rate	Cost
Anxiety	12	0.0211	0.000591	£44.00
Constipation	32	0.0562	0.00161	£44.00
Depression	6	0.0105	0.000294	£44.00
Diarrhoea	28	0.0492	0.00140	£44.00
Dizziness	15	0.0264	0.000741	£44.00
Dry mouth	11	0.0193	0.000542	£44.00
Headache	53	0.0931	0.00271	£44.00
Hot flush	7	0.0123	0.000343	£44.00
Insomnia	29	0.0510	0.00145	£44.00
Nasopharyngitis	31	0.0545	0.00155	£44.00
Nausea	30	0.0527	0.00150	£44.00
Sinusitis	34	0.0598	0.00171	£44.00
Upper respiratory tract infection	64	0.113	0.00331	£44.00
Vomiting	14	0.0246	0.000691	£44.00
Total adverse event cost per week:				£0.81

The level of reporting of AE data across the orlistat studies identified in Section 4.10 and EMA regulatory documents was not sufficient to compare AE incidence accurately to NB32 patients. Based on the expected non-inferior safety profile of NB32 versus orlistat, the model conservatively assumes the same cost per week for patients treated with orlistat as patients treated with NB32.

Total AE costs per week are presented in Table 71.

Table 71: Total adverse event costs

Treatment	Adverse event cost (per week)	Source
NB32	£1.69	Pooled Phase III COR studies, NHS reference costs (2015)
SM	£0.81	
ORL	£1.69	As per NB32
Key: COR, Contrave obesity research; NB, naltrexone 32mg plus bupropion; NHS, National Health Service; ORL, orlistat; SM, standard management.		

5.5.5 Miscellaneous unit costs and resource use

No miscellaneous unit costs or resource use identified.

5.6 Summary of base case de novo analysis inputs and assumptions

5.6.1 Summary of base case *de novo* analysis inputs

A summary of the base case model inputs is presented in Appendix 18 (Table 28).

5.6.2 Assumptions

The main assumptions attributable to modelling methodology are presented within Table 53 in Section 5.2.2.4. All other modelling assumptions relating to input data are described through Sections 5.3, 5.4 and 5.5.

5.7 Base case results

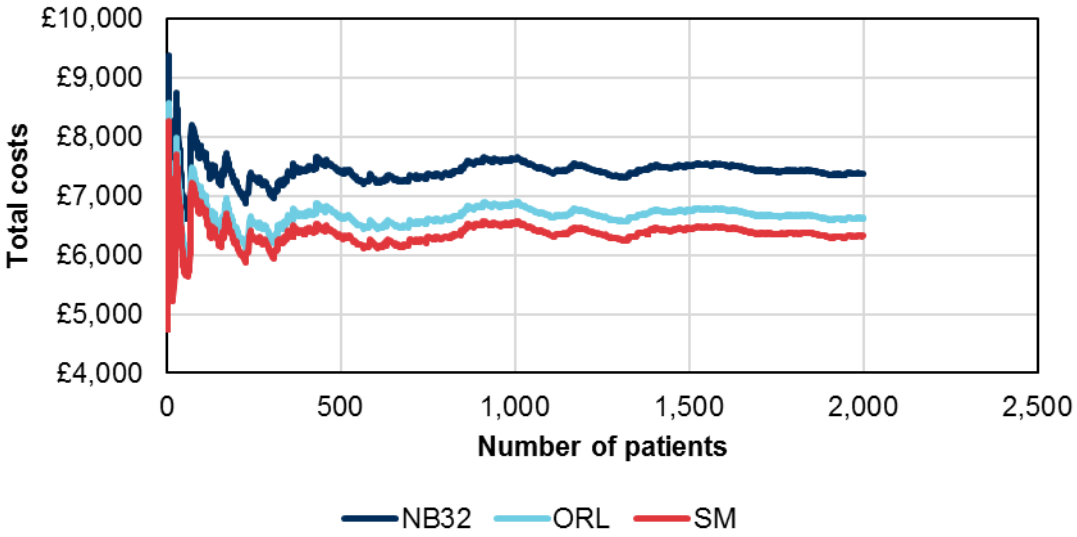
To produce base case results from the DICE model, a sufficient number of patients (or “patient profiles”) are required to be run such that the model results converge to a

consistent value. To establish how many patient profiles are required to produce stable model results, a diagnostic exercise was carried out.

To undertake the diagnostic exercise, the model was run for a maximum of 2,000 randomly-sampled patient profiles. The moving average of the total costs and total QALYs was recorded, which were subsequently plotted on a figure to illustrate how many model runs are required for results to stabilise.

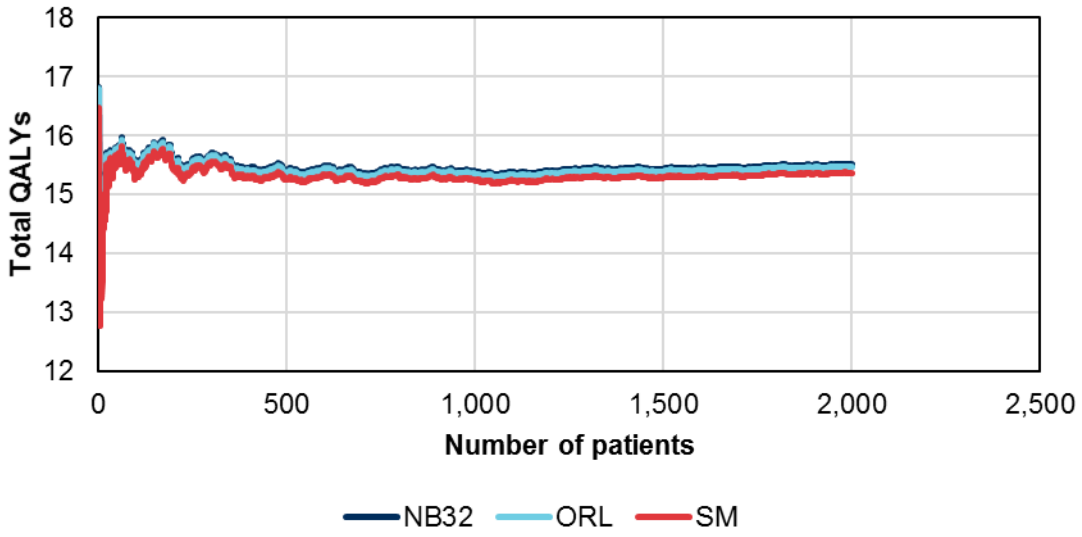
These results of the diagnostic exercise are shown for the total costs and total QALYs in Figure 34 and Figure 35, respectively.

Figure 34: Diagnostic exercise – total costs



Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

Figure 35: Diagnostic exercise – total QALYs



Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

The results of this diagnostic exercise suggested that total estimated QALYs begin to stabilise after results have been collected for approximately 500 simulated patients (Figure 35). However, for total costs the number of patient profiles required to produce stable results was larger (Figure 34). Therefore, 1,000 patient profiles were deemed appropriate for eliciting deterministic model results with an appropriate level of precision while also considering run time.

For probabilistic sensitivity analysis (PSA), a trade-off between the number of patient profiles and the number of probabilistic draws was made. The smallest number of patient profiles required after which model results appear to stabilise may be considered at approximately 500, after which large amounts of model variation do not appear to influence results greatly. Therefore, within the PSA, 500 patient profiles are used for each PSA run. The number of PSA runs was chosen at 100 again, as a direct result of the run time required.

5.7.1 Base case incremental cost effectiveness analysis results

Base case incremental cost-effectiveness analysis results are shown in Table 72.

NB32 adjunct therapy is estimated to offer an additional 0.0765 QALYs per patient versus standard management alone, and an additional 0.0192 QALYs per patient

versus orlistat adjunct therapy. These QALY gains are estimated to cost an incremental £1,044 versus standard management and £750 versus orlistat. Thus, NB32 is estimated to be a cost-effective alternative for NHS patients currently receiving standard management alone, with an incremental cost-effectiveness ratio (ICER) of £13,647 per QALY gained. The estimated ICER versus orlistat adjunct therapy is higher, at £32,084 per QALY gained.

As stressed in Section 1, and illustrated throughout Sections 5.1 to 5.5, conservative assumptions were made in both comparisons. Most notably, the obesity-related health conditions the analysis considers are limited to MI, stroke and T2DM. Section 5.4.2 listed several further weight-related comorbidities in which HRQL data were identified. European Guidelines for Obesity Management in Adults published in 2015 list a total of 63 obesity-related health risks and complications.[6] The blindness of the analysis to many cost and health benefits of weight loss means that the cost-effectiveness of more effective alternatives is inherently underestimated. The results in Table 72 should be interpreted accordingly.

Model estimates for orlistat adjunct therapy are further limited by the key assumptions required to estimate the relative effectiveness of orlistat versus NB32 or standard management alone, and treatment duration for orlistat patients. The need for these assumptions, outlined in Section 5.3 and discussed further in Section 5.11, adds important uncertainty to the conservative comparison to orlistat that the model cannot address. The estimated ICER for NB32 versus orlistat should be interpreted with particular caution; the true ICER could well imply NB32 is a cost-effective alternative to orlistat adjunct therapy, but it is beyond the capability of the economic analysis to demonstrate this.

Table 72: Base case results

T	Total			Incremental			ICER (QALYs)	
	Costs	LYs[a]	QALYs	Costs	LYs[a]	QALYs	Versus baseline (SM)	Incremental
SM	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4148	£294	0.0383	0.0531	£5,538	£5,538
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£13,647	£32,084

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.

Table 73: Pairwise cost-effectiveness results: NB32 versus standard management

T	Total costs	Total LYs[a]	Total QALYs	Incremental costs	Incremental LYs[a]	Incremental QALYs	ICER (QALYs)
SM	£6,519	33.4768	15.3616				
NB32	£7,563	33.5343	15.4381	£1,044	0.0575	0.0765	£13,647

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

Table 74: Pairwise cost-effectiveness results: NB32 versus orlistat

T	Total costs	Total LYs[a]	Total QALYs	Incremental costs	Incremental LYs[a]	Incremental QALYs	ICER (QALYs)
ORL	£6,814	33.5151	15.4148				
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£32,084

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

5.7.2 Clinical outcomes from the model

Clinical outcomes from the model in terms of LYs and QALYs are presented within Section 5.7.1. In addition to these outcomes, weight loss outcomes produced by the model were compared with those used to inform the model as input data. The results of this comparison are shown in Table 75.

Table 75: Clinical weight loss outcomes

Technologies	Outcome	Input data	Model
NB32	Weight loss at primary assessment Week 16 (responders)	9.4%	9.4%
	Weight loss at secondary assessment Week 56 (responders)	11.7%	12.0%
ORL	Weight loss at primary assessment Week 12 (responders)	8.6%	8.8%
	Weight loss at secondary assessment Week 52 (responders)	10.9%	11.2%
SM	Weight loss at primary assessment Week 12	2.3%	2.2%
	Weight loss at secondary assessment Week 52	4.5%	4.5%

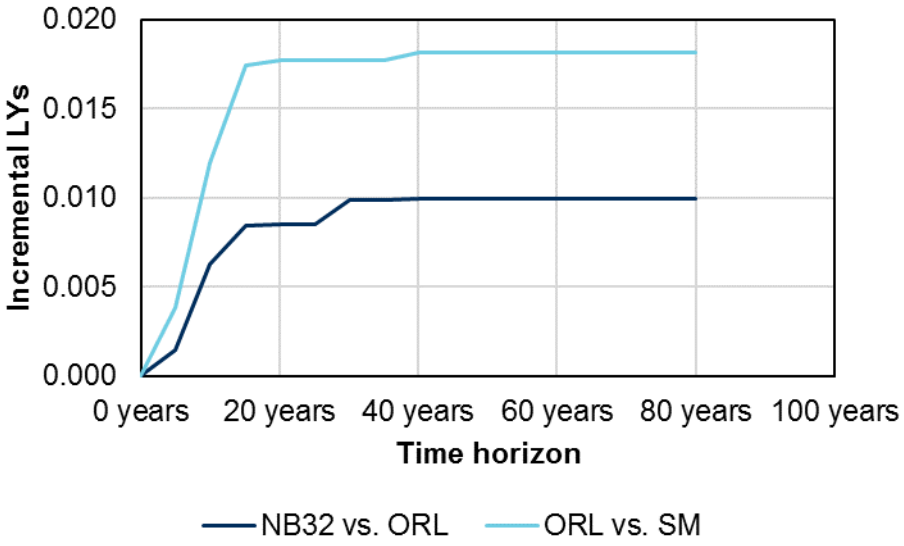
Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

The results show that the values produced by the model are within an acceptable range of those used to inform the model, given the variability associated with sampling a large range of random numbers and 1,000 patient simulations conducted.

In addition to the weight loss-based outcomes shown in Table 75, Figure 36 and Figure 37 show the incremental LYs and QALYs accrued by patients over time, respectively. The analysis was conducted using a sampled profile of 500 patients, as after this number of patients, the QALY outcome was shown to stabilise within the diagnostic exercise presented in Figure 35.

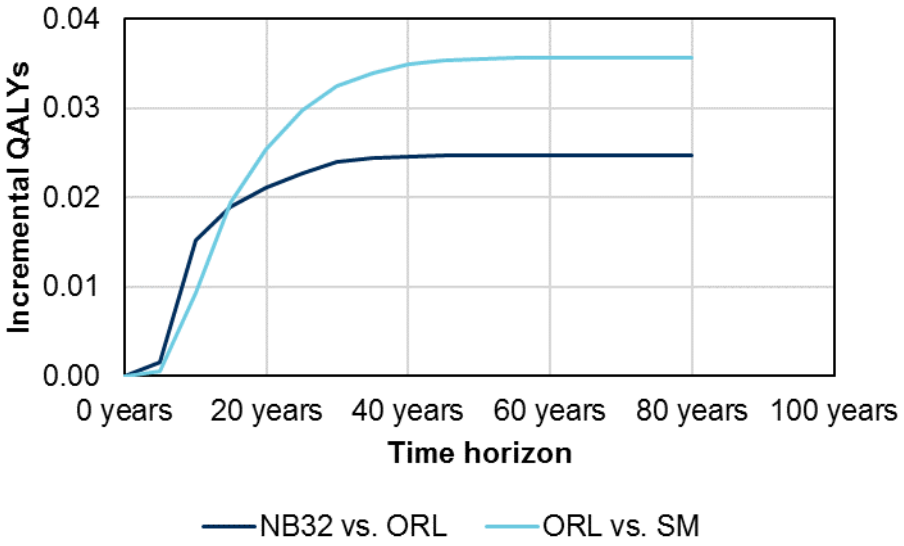
As expected, LYs are shown to increase consistently over time, with a plateau shown when most patients have reached an age in line with their life expectancy. The UK life expectancy at the starting age for the sampled cohort is expected to be between approximately 34 and 37 years (for males and females, respectively) in the general population.[118] For QALYs, a similar pattern is demonstrated.

Figure 36: LYs accrued over time



Key: LY, life year; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

Figure 37: QALYs accrued over time



Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

5.7.3 Disaggregated results of the base case incremental cost effectiveness analysis

Table 76 shows the discounted total costs incurred by patients over the modelled time horizon across all treatment arms, separated by cost category. Within the DICE model, costs were assigned to one of the following categories:

- Treatment acquisition – the cost of NB32 or orlistat. For standard management patients, this cost is £0.
- Standard management and condition management – the cost of non-pharmacological standard management (i.e. GP visits, nurse visits and blood tests), as well as costs associated with T2DM, MI and stroke.
- AEs – all costs relating to the treatment of AEs.
- Death – all costs relating to the cost of cardiovascular-related mortality.

Table 76: Summary of discounted costs by cost category

Technologies	Costs				
	Treatment acquisition	SM and CM	AEs	Death	Total
SM	£0	£5,982	£171	£367	£6,519
ORL	£238	£5,993	£216	£366	£6,814
NB32	£995	£5,983	£220	£366	£7,563

Key: AE, adverse event; CM, condition management; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

The results show that the majority of estimated costs relate to the standard management and condition management of patients. Evidence cited in Section 3 estimated that NHS costs attributed to elevated BMI were as high as £15.8 billion over 10 years ago[2], and given the limited extent to which BMI-linked health conditions are captured in the analysis, the per-patient lifetime costs of obesity management in Table 76 are inherent underestimates of the true costs.

Table 76 shows total standard management and condition management to be similar across treatment arms. This succinctly highlights how conservative the analysis is. If the analysis was informed by (i) evidence on probability of MI, stroke and MI events as well as time to these events, and (ii) the requisite evidence on more of the 63

known obesity-related health events and conditions[6], NB32 patients would be estimated to have a substantially lower standard management and condition management cost burden than similar patients receiving less effective alternatives.

In line with expectations given the assumptions outlined in Section 5.5, the costs of treating AEs are estimated to be broadly similar across treatment arms, as are the estimated costs associated with death.

Table 77 shows the undiscounted total costs incurred by patients across all treatment arms, separated by cost category.

Table 77: Summary of undiscounted costs by cost category

Technologies	Costs				
	Treatment acquisition	SM and condition management	Adverse events	Death	Total
SM	£0	£11,895	£185	£1,065	£13,144
ORL	£247	£11,894	£232	£1,065	£13,438
NB32	£1,034	£11,878	£236	£1,065	£14,212
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.					

Disaggregated results for LYs and QALYs are not available from the model. This is a direct consequence of the chosen model structure, as there are no distinct Markovian health states from which disaggregated LYs and QALYs may be drawn. However, summaries of the LYs and QALYs gained over time in the model are presented within Section 5.7.2.

5.8 Sensitivity analyses

5.8.1 Probabilistic sensitivity analysis

As discussed in Section 5.7, a sampled patient profile of 500 patients was used within each PSA model run and 100 PSA runs were simulated for the 500 patients, with mean results recorded for each iteration.

A comparison of the mean probabilistic base case model results with the deterministic base case model results are shown in Table 78. The probabilistic results are in line with the deterministic results; however, these results are limited in terms of the number of PSA runs and size of the sampled patient profile.

Table 78: Comparison of 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	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4148	£294	0.0383	0.0531	£5,538	£5,538
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£13,647	£32,084
Probabilistic base case model results								
SM	£6,411	33.5673	15.3664					
ORL	£6,667	33.6128	15.4176	£256	0.0455	0.0512	£4,993	£4,993
NB32	£7,409	33.6242	15.4379	£742	0.0115	0.0204	£13,936	£36,405

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.

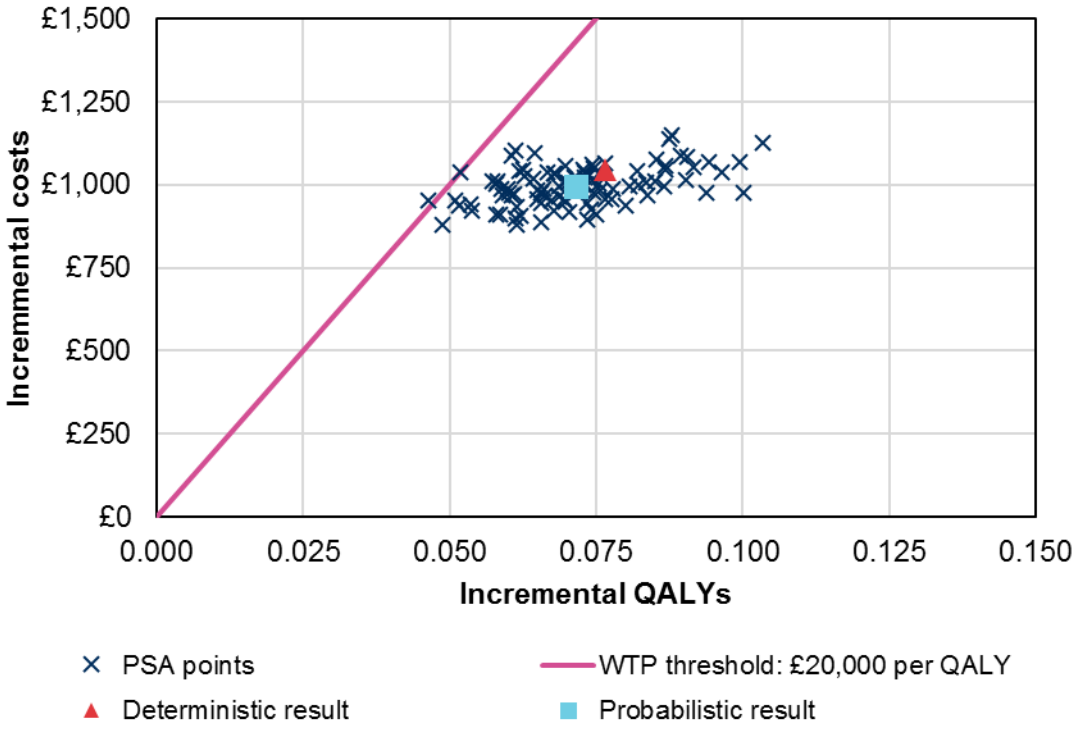
Figure 38 shows the PSA scatterplot for NB32 versus standard management. The scatterplot demonstrates some parameter uncertainty around the mean model result. However, all probabilistic model runs appear to demonstrate results that are not dissimilar to the probabilistic and deterministic mean results.

Importantly, it was not possible to incorporate parameter uncertainty around natural history model parameter estimates into sensitivity analyses, owing to reporting in Ara et al.[14] As such, the PSA is unable to fully demonstrate the consequence of parameter uncertainty for uncertainty around key model results.

In addition, much of the key uncertainty around model results is structural and methodological, 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.

The PSA requires an external datafile to inform the random number draws used in the model equations. Details of how the PSA was carried out within the model are presented in Appendix 19.

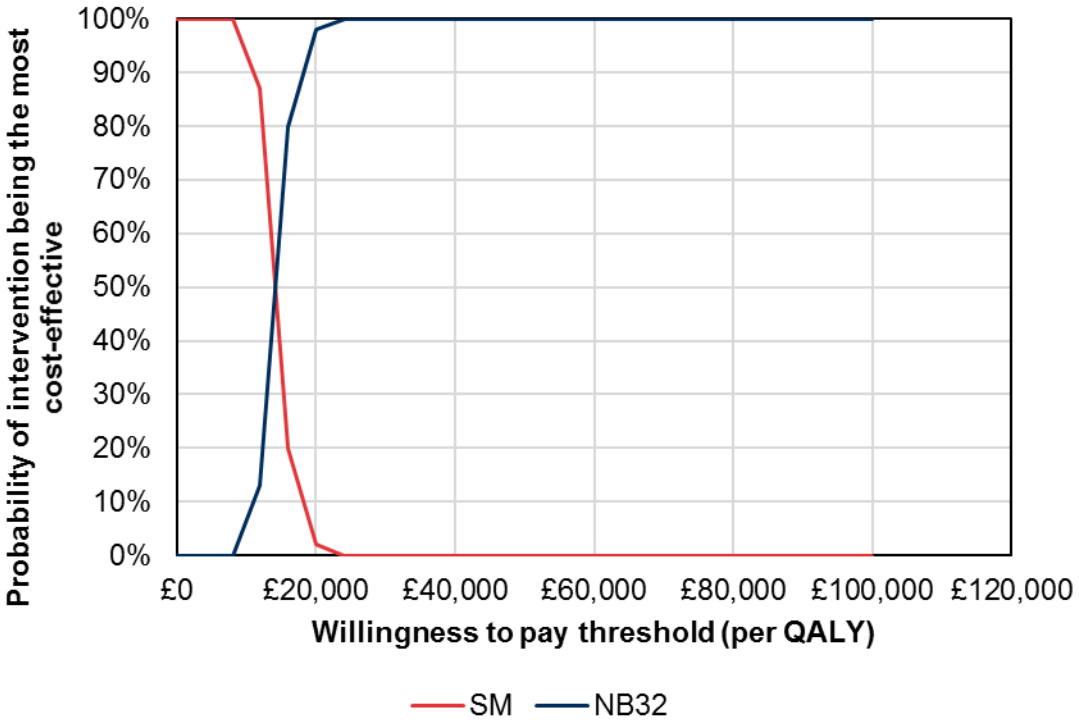
Figure 38: PSA scatterplot – NB32 versus SM



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

Figure 39 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 a 98% probability of being cost effective versus standard management at a willingness to pay (WTP) threshold of £20,000 per QALY gained.

Figure 39: CEAC – NB32 versus SM

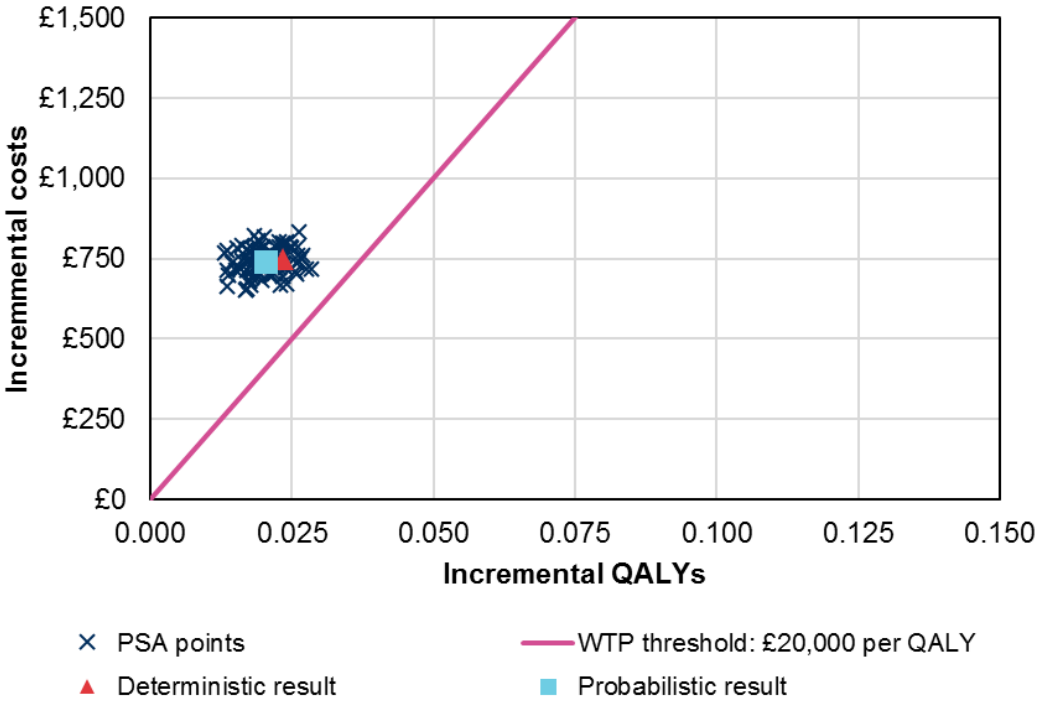


Key: NB32, naltrexone 32mg plus bupropion; PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life year; SM, standard management.

Figure 40 shows the PSA scatterplot for 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. Figure 41 shows the CEAC for NB32 versus orlistat. The CEAC suggests that NB32 is associated with a 0% probability of being cost effective versus orlistat at a WTP threshold of £20,000 per QALY gained.

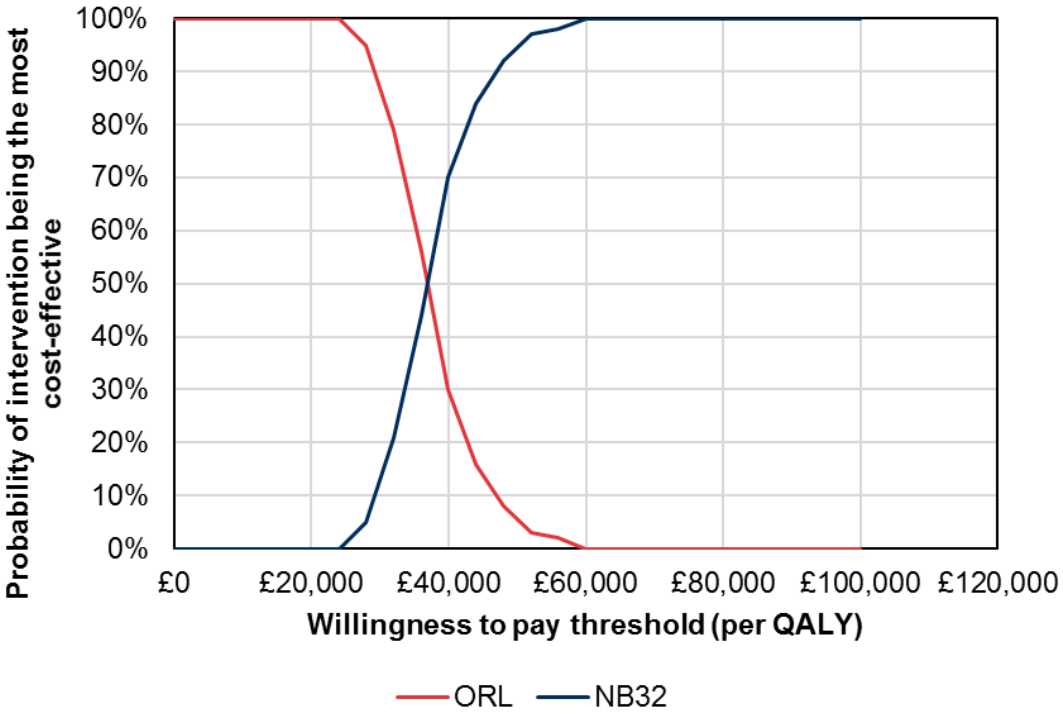
As stressed with respect to the comparison to standard management alone, PSA results for the comparison to orlistat should be interpreted with care. Much of the key uncertainty around model results is structural or methodological, and based in the key conservative assumptions underpinning the analysis. The true probability that NB32 is a cost-effective alternative to orlistat is not zero. It is highly plausible that the estimated probability that NB32 is preferable to orlistat would be greater than 50% 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 40: PSA scatterplot – NB32 versus ORL



Key: NB32, naltrexone 32mg plus bupropion; PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life year; ORL, orlistat; WTP, willingness to pay.

Figure 41: CEAC – NB32 versus ORL



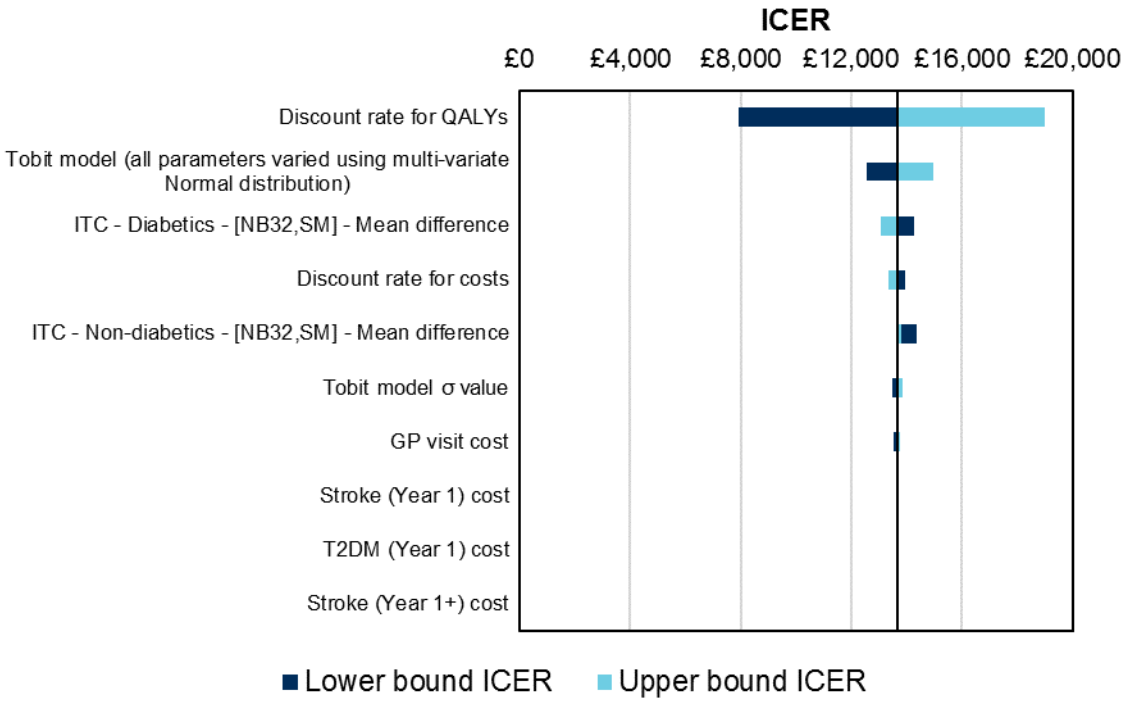
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; PSA, probabilistic sensitivity analysis; QALY, quality-adjusted life year.

5.8.2 Deterministic sensitivity analysis

One-way sensitivity analysis (OWSA) was conducted to quantify the uncertainty associated with each model parameter in the model results. Within OWSA, all relevant model parameters were varied between their lower and upper bounds, and the model result was recorded.

Figure 42 presents the 10 most influential parameters on model results for NB32 versus standard management in the form of a tornado diagram.

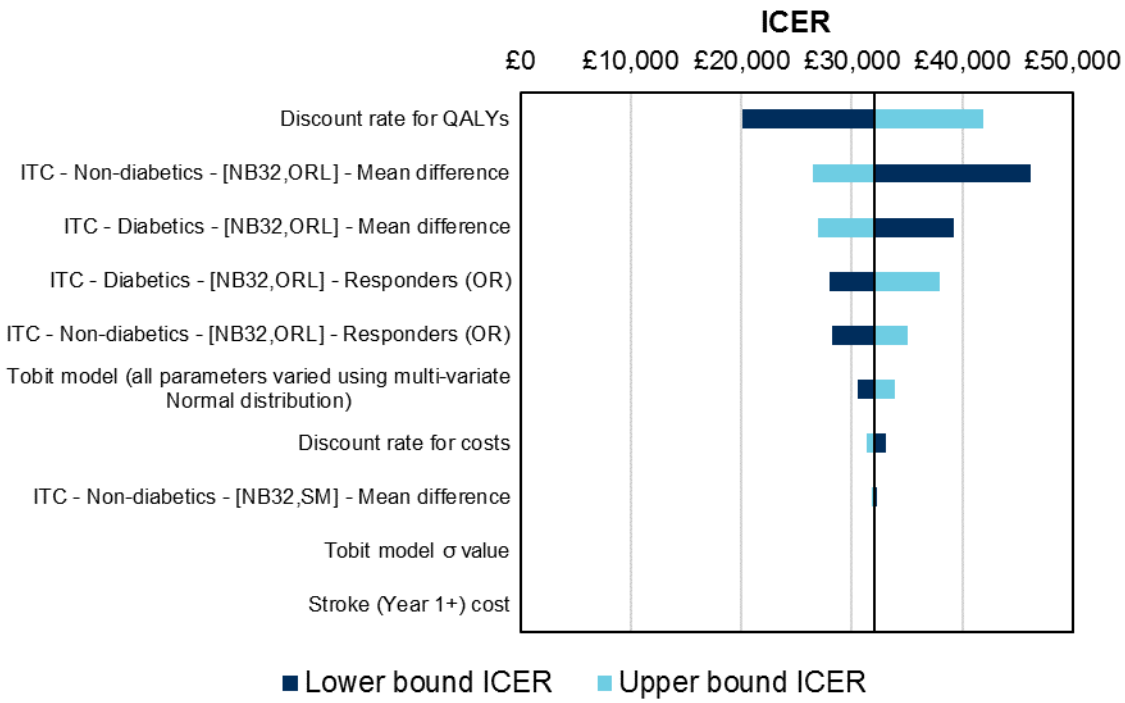
Figure 42: OWSA – NB32 versus SM



Key: GP, general practitioner; ICER, incremental cost-effectiveness ratio; ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; OLS, ordinary least squares; QALY, quality-adjusted life year; SM, standard management; T2DM, Type 2 diabetes mellitus.

Figure 43 presents the 10 most influential parameters on model results for NB32 versus orlistat in the form of a tornado diagram.

Figure 43: OWSA – NB32 versus ORL



Key: GP, general practitioner; ICER, incremental cost-effectiveness ratio; ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; OLS, ordinary least squares; OR, odds ratio; QALY, quality-adjusted life year; SM, standard management; T2DM, Type 2 diabetes mellitus.

Note: The eighth parameter (ITC – Non-diabetics – [NB32, SM] – Mean difference] is not an error. This parameter is featured within the outcome of the analysis as patients who discontinue treatment with orlistat may continue treatment with standard management alone.

The most influential parameters on model results are those relating to the HRQL of patients (i.e. the Tobit model and the discount rate for QALYs), as well as those related to the measures of relative efficacy from the ITC. All other model parameters have a negligible impact on model results.

For the comparison of NB32 with standard management, no parameter was shown to produce an ICER of more than £20,000 per QALY gained.

For the comparison of NB32 with orlistat, relatively large amounts of variation were shown, largely in line with the uncertainty attributable to the ITC. All comparisons to orlistat are severely limited due to the availability of data and the lack of directly comparative evidence.

In addition, as described in Section 5.8.1, much of the key uncertainty around model results is structural or methodological, as opposed to the parameters explored within OWSA. Therefore, although results in this analysis pertain to quantifiable uncertainty regarding the cost effectiveness of NB32 versus orlistat, the results should be interpreted in consideration of the evidence available, and the limitations of the economic analysis used to produce them.

5.8.3 Scenario analysis

Scenario analysis was undertaken to explore the impact of specific scenarios on cost-effectiveness results. The following scenarios were explored:

- The time period over which weight is regained:
  - Weight regain is set at 3 years in the base case model results. Scenario analysis was undertaken to explore the impact on results if this value were varied largely between 2 years and 5 years.
- The cost of T2DM
  - The cost of T2DM was taken from an alternative source, as opposed to being lifted directly from the report by Ara et al., as the cost presented in their report did not consider any ongoing costs for the treatment of T2DM beyond 1 year.[14] However, use of the cost from the report by Ara et al. (inflated using HCHS inflation indices) was explored as a scenario analysis.[179]
- Structural assumptions implicit in the HSE EQ-5D data analysis informing utility assumptions
  - The OLS regression results from the Copley et al., presented in Section 5.4.2 alongside base case Tobit model results, are used as a structural alternative to the Tobit model estimates.
- The cost of AEs
  - The cost of treating AEs is set within the model base case as a visit to the GP. As a scenario analysis, the cost of treating all AEs were assumed to be the cost of an outpatient consultation.

- Discounting
  - NICE guidance states that where health benefits are sustained over a very long period (normally at least 30 years), the Appraisal Committee may apply discount rates of 1.5% for health effects and 3.5% for costs. Hence, as a scenario analysis, discount rates of 3.5%, 1.5% and 0% were applied for costs, QALYs and LYs, respectively.[115]
- Time horizon
  - A time horizon of 15 years was analysed as a scenario analysis to ascertain cost-effectiveness estimates within a shorter time horizon than lifetime. This time horizon was selected to align with the limitations of the GPRD data informing the BMI natural history and TTE models underpinning the analysis.

Table 79 contains the results of scenario analysis undertaken on key areas of uncertainty within the model. The results show that the most influential scenarios on model results were those relating to the time horizon over which costs are incurred and benefits are accrued. Also of consequence were assumptions around discount rates and the time over which weight is expected to be regained.

Given the small estimated differences in mean patient costs and outcomes across treatments in the base case analysis, and the resulting sensitivity of the ICER as a measure of outcome, the results are robust to changes to many key assumptions. The ICER versus standard management is below £15,200 in all scenarios bar the scenario in which the time horizon is restricted to 15 years. In this scenario, the ability of the analysis to capture health and cost benefits of delays in TTE is severely curtailed, and therefore, the cost effectiveness of an effective treatment is underestimated even more greatly than in the base case. Although it is important to illustrate and explain the sensitivity of results to time horizon assumptions, this scenario should not be used to inform decision making.

Table 79: Scenario analysis results

Scenario				ICERs	
				NB32 vs	
n	**Model setting**	**Base case**	**Scenario tested**	**ORL**	**SM**
0	Base case			£32,084	£13,647
1	Weight regain	3 years	2 years	£41,016	£14,113
2	Weight regain	3 years	5 years	£29,739	£11,880
3	Cost of T2DM	£347.57	£175.86 in Year 1 only	£36,096	£13,764
4	Utility model	Tobit	OLS	£36,771	£10,285
5	AE costs	All GP	All outpatient	£36,492	£15,130
6	Discounting	3.5% for costs & effects	1.5% for costs & effects	£28,323	£9,969
7	Time horizon	Lifetime	15 years	£53,514	£22,763

Key: AE, adverse event; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; NB32, naltrexone 32mg plus bupropion; OLS, ordinary least squares; QALY, quality-adjusted life year; SM, standard management; T2DM, Type 2 diabetes mellitus.

5.8.4 Summary of sensitivity analyses results

The sensitivity and scenario analyses presented in this section were designed to capture the uncertainty around results that stems from uncertainty around model inputs and assumptions, where possible.

The key areas of uncertainty highlighted by the sensitivity analyses related to the HRQL of patients (Section 5.4.4) and the rate of weight regain (Section 5.2.2.4).

PSA results suggest that NB32 is a high cost-effective treatment option in combination with standard management compared with standard management alone, with a probability of 98% that the ICER lies below £20,000 per QALY gained. NB32 was also shown to remain an effective treatment option when compared with orlistat.

The PSA conducted within the model is limited by the availability of data to explore the uncertainty of the equations that inform the model (i.e. the risk equations). However, as previously discussed, the model itself provides conservative cost-effectiveness estimates as these risk equations are concerned solely with the anticipated time to a given clinical event (as opposed to the probability of experiencing such an event), and therefore, although PSA does not consider these

equations within the analysis, the key uncertainty relating to the model is structural in nature and is therefore not captured within the sensitivity analysis.

OWSA demonstrated that the model is most sensitive to inputs relating to the HRQL of patients (i.e. the Tobit utility regression model and the discount rate for QALYs) and parameters relating to the relative efficacy of treatments. All other model parameters had a negligible impact on model results. Each of the parameters varied within OWSA produced an ICER for NB32 versus standard management of less than £20,000 per QALY gained.

Scenario analyses demonstrated the robustness of the cost-effectiveness estimate for NB32 in combination with standard management versus standard management alone, with ICERs between £9,969 and £22,763 per QALY gained. The most influential scenarios were those relating to the time horizon over which costs are incurred and benefits are accrued, as well as discount rates and the time period over which weight is expected to be regained. The cost-effectiveness estimates for NB32 versus orlistat were also shown to remain within a close range of the deterministic base case results.

Nearly all sensitivity analyses conducted demonstrated a cost per QALY gained for NB32 versus standard management of less than £20,000, with only 2% of PSA runs producing a cost per QALY above £20,000 and one scenario considering a shorter time horizon producing an ICER just over £22,763. The time horizon considered within this analysis should be considered inappropriate within the context of decision making, as it does not allow for the analysis to capture health and cost benefits of delays in the times to events.

Importantly, key conservative assumptions implicit in the analysis have not been explored in the sensitivity analysis. These include the downstream cost and health benefits of effective weight loss therapy for both *time to* and *probability of* all of the 63 obesity-related health risks and complications listed in 2015 European Guidelines for Obesity Management in Adults[6], bar T2DM, MI and stroke, for which only *time to* event risks are included. Even if estimated incremental costs for the comparison between NB32 and orlistat are assumed to be correct, if 0.0142 incremental QALYs are being masked by the key conservative analysis assumptions outlined in this section, the true ICER for NB32 versus orlistat is below £20,000 per QALY gained. If

risk, cost and utility data on just some of the downstream weight-related health benefits of weight loss not captured in the model could be identified and incorporated into the analysis, NB32 could well be shown to dominate both orlistat and standard management in an incremental economic analysis.

5.9 Subgroup analysis

Section 4.10 presents analyses of key clinical endpoints for T2DM and non-T2DM patients separately. Therefore, the model was run using a profile of patients with T2DM at baseline, and again with a profile of patients without T2DM at baseline.

The results of these subgroup analyses are shown in Table 80 and Table 81 for T2DM patients and non-T2DM patients at baseline, respectively.

Table 80: Base case results – T2DM patients at baseline only

T	Total			Incremental			ICER (QALYs)	
	Costs	LYs[a]	QALYs	Costs	LYs[a]	QALYs	Versus baseline (SM)	Incremental
SM	£10,199	32.7296	14.3707					
ORL	£10,496	32.7583	14.4295	£297	0.0287	0.0588	£5,059	£5,059
NB32	£11,216	32.7656	14.4395	£720	0.0073	0.0100	£14,797	£72,069

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

Table 81: Base case results – non-T2DM patients at baseline only

T	Total			Incremental			ICER (QALYs)	
	Costs	LYs[a]	QALYs	Costs	LYs[a]	QALYs	Versus baseline (SM)	Incremental
SM	£3,844	33.5497	15.7335					
ORL	£4,077	33.5854	15.7706	£233	0.0356	0.0371	£6,283	£6,283
NB32	£4,811	33.5944	15.7966	£734	0.0090	0.0259	£15,339	£28,291

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; 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.

In consideration of these results, it should be noted that the ICER is highly sensitive to changes in incremental costs and QALYs, given the relatively small incremental costs and benefits associated with NB32 treatment.

Furthermore, it should be noted that all comparisons to orlistat should be interpreted with care, as data regarding comparisons of NB32 to orlistat in only patients with T2DM are extremely limited as shown, in Section 4.10.

Results for NB32 versus standard management in these subgroups are broadly in line with those produced in the model base case (i.e. assuming 33.2% of patients with T2DM at baseline).

Results for NB32 versus orlistat in these subgroups show a larger range of variability, particularly for the T2DM patients at baseline subgroup. As previously discussed, this comparison is extremely limited due to the data available to compare the cost effectiveness of NB32 and orlistat for patients presenting with T2DM at baseline.

5.10 Validation

5.10.1 Validation of *de novo* cost-effectiveness analysis

External validity

Advice from Professor John Wilding was crucial in informing and validating key clinical assumptions in the analysis. Key input was provided during a 90-minute discussion on 29 September 2016. The notes from this meeting are disclosed as part of this submission, in the interest of transparency.[31] We are grateful to Professor Wilding for his advice at this meeting, and for his openness to further questions up to submission.

The model produces total LYs in the range of 33.48 to 33.53. These values exhibit face validity, given that average age upon entry to the model is approximately 47.0 years, and that UK life expectancy for the general population at this age suggests additional LYs of between approximately 34 and 37 years (for males and females, respectively) in the general population.[118] Furthermore, total QALYs from the *de novo* model are similar to those reported by Ara et al., as Table 82 illustrates.[14]

Table 82: Comparison of Total QALY estimates across the *de novo* analysis and Ara et al.[14]

Technologies	Total Discounted QALYs	
	***De novo* model results**	**Ara et al. results[14]**
SM	15.3616	15.13
ORL	15.4148	15.30
Key: ORL, orlistat; SM, standard management.		

Internal validity

The model was quality-assured by the internal processes of the external economists who adapted the economic model. In this process, an economist not involved in model adaptation reviewed the model for coding errors, inconsistencies and the plausibility of inputs. This included the model being put through a checklist of known modelling errors, and questioning of the assumptions.

5.11 Interpretation and conclusions of economic evidence

The economic analysis has taken a robust and conservative approach to estimate the cost effectiveness of NB32 adjunct therapy for NHS England patients. The approach is consistent with a previous high-quality NIHR-funded systematic analysis of competing drug treatments for overweight and obese patients.[14] The analysis clearly demonstrates NB32 to be a cost-effective adjunct to standard management for patients who would otherwise receive standard management alone.

A key strength of the economic analysis is its methodological robustness. The individual-level, continuous-time approach is advantageous both for its sensitivity to the complexities of the disease area and its suitability for the key natural history data from Ara et al.[14] Section 5.10.1 demonstrates the consistency of model outputs across the *de novo* model and Ara et al.[14] This, and the care taken to ensure assumptions are reflective of NHS practice with key and transparent input from Professor John Wilding, should assure the reader that analysis is designed to reflect clinical practice in England to the limits of practical possibility.

There should be little doubt that base case cost-effectiveness estimates are inherently conservative, and should be interpreted as such. First, the analysis is blind

to cost and HRQL benefits of weight reduction in obese and overweight patients for known risks associated with possibly over sixty health events[6], including numerous cancers[160, 173], hypertension and hyperlipidaemia[150], joint and spinal complaints[31, 126, 128, 134], multiple sclerosis[152], and sleep apnoea.[31] As such, the cost effectiveness of the treatment with the greatest effectiveness in terms of weight reduction, NB32 adjunct therapy, is inherently underestimated.

In addition, while the natural history risk models capture the effect of weight reduction upon *time to* T2DM onset, MI and stroke based on large-scale UK patient data, the effect of weight reduction upon *probability of* T2DM onset, MI and stroke is not captured. This is another important and inherently conservative assumption in the *de novo* model inherited from Ara et al.[14], and should be considered when interpreting results.

A third key conservative feature of the analysis are assumptions around treatment discontinuation. Weight regain is assumed to begin upon treatment discontinuation. In addition, treatment is assumed to end at the limit of clinical trial data. Again, this is consistent with Ara et al.[14], but it underestimates the economic value of NB32 if patients continue to benefit from effective weight reduction treatment after discontinuation, or if a proportion of patients continue treatment beyond the point where they are lost to follow-up in clinical trials.

The base case ICER versus standard management was shown in Section 5.7 to be around £13,600 per QALY gained, and the robustness of this estimate to testable parameter and structural uncertainty explorations was shown in Section 5.8. These findings are testament to the clear value of NB32 adjunct therapy for NHS patients who would otherwise receive only standard non-pharmacological management.

Prior treatment with orlistat adjunct therapy, a drug with a totally different mechanism of action to NB32, should have no impact on the effectiveness of NB32 adjunct treatment. As such, NB32 offers a further pharmacological treatment option to patients who have failed to achieve adequate weight loss with orlistat treatment, or who failed to comply with dietary requirements associated with orlistat, or were unable to tolerate orlistat treatment and would otherwise revisit standard management measures.

The main weakness of the analysis is its limited ability to provide an accurate economic comparison to orlistat adjunct therapy. This is a direct consequence of the regulatory treatment discontinuation rule that applies in clinical practice but was not used in the key clinical trials. Therefore, summary comparative effectiveness estimates from orlistat RCT publications are not reflective of clinical practice beyond 12 weeks. This was a challenge for economic appraisal of NB32 adjunct therapy versus standard management too. However, for NB32, this could be addressed with analysis of patient-level data from the COR trial programme and NB-CVOT study. No such orlistat trial patient data were available to the company. Consequently, relative effectiveness estimates from the NMA, described in Section 4.10 and informing the model as described in Section 5.3.3, do not account for regulatory response-based treatment stopping rules. Further substantial orlistat treatment discontinuation assumptions were required in the *de novo* model in the absence of data.

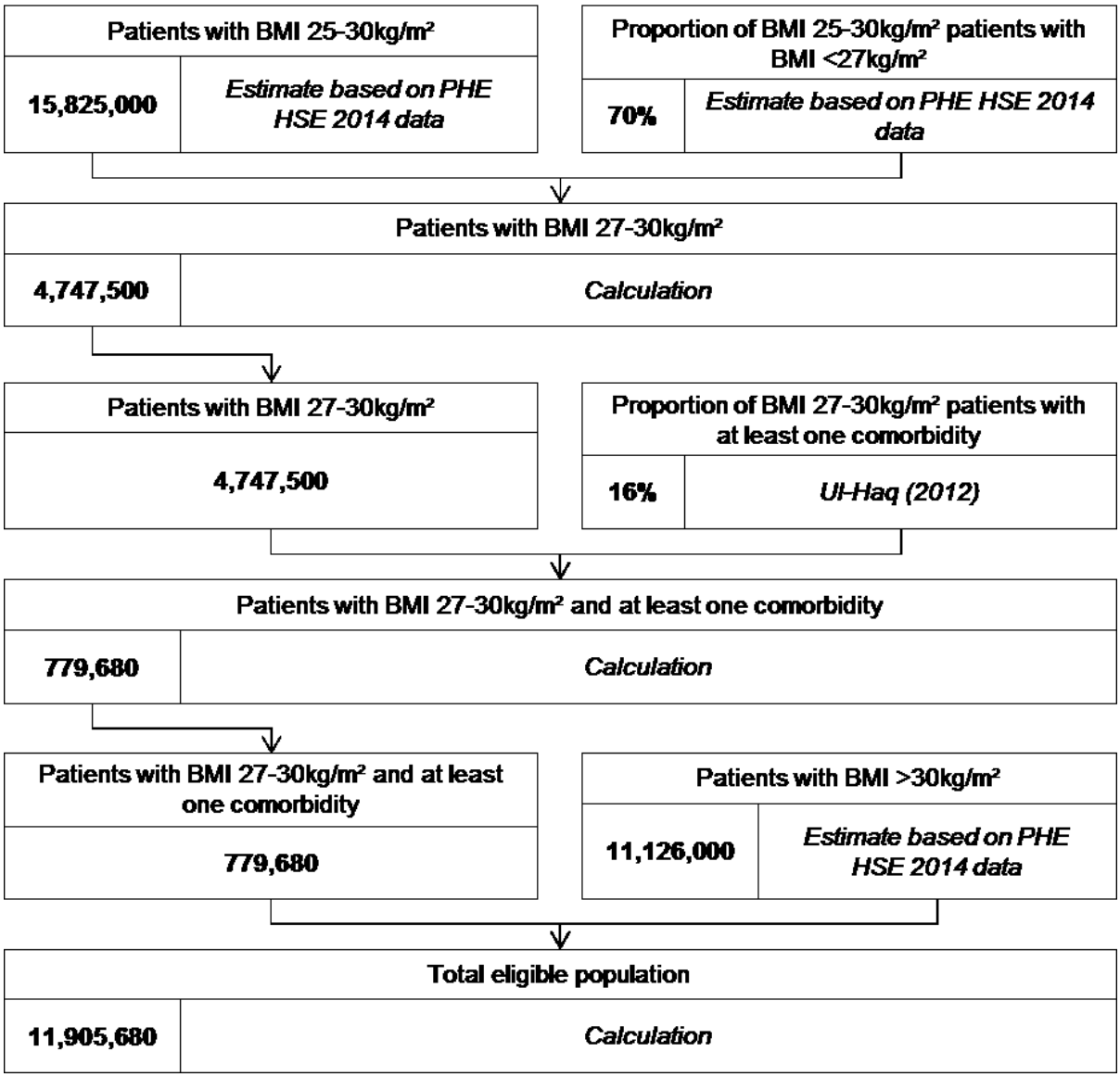
The base case ICER versus orlistat adjunct therapy, viewed in isolation, suggests that NB32 adjunct therapy for patients who would otherwise receive orlistat adjunct therapy would not be cost effective at the NICE WTP threshold. However, this estimate should be interpreted with caution. The estimated patient QALY benefit and incremental cost of NB32 versus orlistat are small (0.0234 QALYs and £750). Given the conservative features of the *de novo* model outlined in this section, in this comparison, it is very likely that the true incremental costs of NB32 have been overestimated, while the true incremental benefits were underestimated.

As incremental cost and QALY estimates for this comparison are small, the ICER is sensitive. Even if estimated incremental costs are assumed to be correct, if 0.0142 incremental QALYs are being masked by the key conservative analysis assumptions outlined in this section, the true ICER for NB32 versus orlistat is below £20,000 per QALY gained. If the downstream cost and health benefits of effective weight loss therapy for both *time to* and *probability of* even a few more of the 63 obesity-related health risks and complications listed in 2015 European Guidelines for Obesity Management in Adults could be incorporated into the analysis[6], the clear economic value of NB32 for NHS England patients could be far better demonstrated.

6 Assessment of factors relevant to the NHS and other parties

The total number of patients eligible for treatment with NB32 was required to derive budget impact estimates. A range of sources were used to derive these estimates, as shown in Figure 44.

Figure 44: Derivation of eligible patient population



Key: BMI, body mass index; HSE, Health Survey England; PHE, Public Health England.
Sources: PHE HSE (2014)[41]; Ul-Haq (2012)[191]

However, the population of patients currently receiving standard management is much smaller than the eligible population. Based on HSCIC-QOF data, 4,186,000 patients currently receive standard management treatment within the NHS.[192]

Based on the population of patients who currently receive standard management, an overview of the expected eligible patient population was produced over a 5-year period. These figures are presented in Table 83.

Table 83: Total eligible patients within the budget impact analysis

	2017	2018	2019	2020	2021
Total eligible patients	4,275,144	4,320,425	4,366,186	4,412,431	4,459,167

Note: The estimated population was increased over the 5-year period using an estimated obese population annual growth rate of 1.1%, based on Public Health England Obesity Knowledge Information Team data over a 10-year period.

$$Annual\ growth\ rate = \left(\left(\frac{25\%}{22.5\%}\right)^{1/10}\right) - 1 = 1.1\%$$

Following derivation of the number of patients currently receiving standard management, the number of patients currently receiving orlistat was estimated using prescription cost analysis (PCA) data.[178] As data were only available for the number of orlistat packs prescribed, an estimate of the average number of packs per patient was derived from the *de novo* model to elicit expected patient numbers treated with orlistat, which produced an estimate of approximately 13.4 packs per patient. The figures produced using these data are presented in Table 84.

Table 84: Total patients treated with orlistat (2011–2016)

	2011	2012	2013	2014	2015	2016
Number of packs	875,829	570,839	509,751	535,898	496,473	465,296
Estimated patient numbers	65,394	42,622	38,061	40,013	37,069	34,741

Using the average annual change from 2012 to 2016, a reduction in the number of patients treated with orlistat from 2017 onwards of 4.0% per annum was applied. This yielded the expected numbers of orlistat patients shown in Table 85.

Table 85: Total patients treated with orlistat (2017–2021)

	2017	2018	2019	2020	2021
Estimated patient numbers	33,349	32,013	30,731	29,500	28,318

The estimated market share for NB32 was derived using an estimated [redacted] of patients in Year 1 who would have otherwise been treated with orlistat who would now be treated with NB32. In addition, a further [redacted] of this estimated number of NB32 patients (previously orlistat) are expected to also be treated with NB32, but would have previously received standard management alone. These figures were expected to increase by [redacted] per annum.

The total figures are shown in Table 86. Uptake is expected to be approximately [redacted] patients in Year 1, increasing to approximately [redacted] patients by Year 5.

Table 86: Total patients treated with NB32 (2017–2021)

	2017	2018	2019	2020	2021
Patients previously treated with ORL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Patients previously treated with SM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.					

Based on these figures, the following scenarios may be considered:

- **Scenario A:** No use of NB32 (i.e. current overview)
- **Scenario B:** NB32 is introduced in place of standard management treatment
- **Scenario C:** NB32 is introduced in place of both standard management treatment and orlistat treatment

The patient numbers shown in Table 87 were used to inform budget impact calculations.

Table 87: Patient numbers for all budget impact scenarios

Scenario	Treatment	2017	2018	2019	2020	2021
Scenario A	NB32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ORL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	SM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario B	NB32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ORL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	SM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Scenario C	NB32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	ORL	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	SM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total		[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: NB32, naltrexone 32mg plus bupropion; NHS, National Health Service.

To produce budget impact estimates, the *de novo* model was run with restricted time horizons of 1 to 5 years. The total costs for each treatment arm were recorded and used to inform the expected costs for all patients incurred in each calendar year.

As data from the *de novo* model were used to inform budget impact estimates, only treatment acquisition costs were considered within the analysis. This was considered appropriate as the long-term benefits of treatment with NB32 are not sufficiently captured within a 5-year time horizon, and should therefore not be considered in isolation of the downstream costs and benefits of treatment with NB32.

Furthermore, the *de novo* model does not fully illustrate the downstream costs and benefits of weight reduction, as the modelled equations relate to the predicted time to a clinical event, rather than the probability of experiencing an event. Consequently, as the full potential benefits of weight loss were not captured within the *de novo* model, the budget impact estimates presented here may be over-predictive of the true budget impact of NB32. In consideration of a broader perspective, NB32 treatment could even lead to cost savings from an NHS perspective.

The total budget impact figures for Scenarios A, B and C are shown in Table 88, Table 89 and Table 90, respectively.

Table 88: Scenario A: Budget impact results

		2017	2018	2019	2020	2021
NB32	TA	£0	£0	£0	£0	£0
ORL	TA	£4,411,626	£6,756,739	£7,788,895	£7,476,855	£7,177,317
SM	TA	£0	£0	£0	£0	£0
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management; TA, treatment acquisition.						

Table 89: Scenario B: Budget impact results

		2017	2018	2019	2020	2021
NB32	TA	£436,073	£926,872	£1,543,162	£2,314,768	£3,472,152
ORL	TA	£4,411,626	£6,756,739	£7,788,895	£7,476,855	£7,177,317
SM	TA	£0	£0	£0	£0	£0
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management; TA, treatment acquisition.						

Table 90: Scenario C: Budget impact results

		2017	2018	2019	2020	2021
NB32	TA	£2,180,367	£4,634,360	£7,715,811	£11,573,841	£17,360,762
ORL	TA	£3,970,464	£5,842,810	£6,287,717	£5,225,088	£3,799,666
SM	TA	£0	£0	£0	£0	£0
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management; TA, treatment acquisition.						

Incremental budget impact results for treatment acquisition costs and all costs are shown for Scenario A (no NB32 use) versus Scenario B (displacement of standard management) and versus Scenario C (displacement of both standard management and orlistat) in Figure 45 and Figure 46, respectively.

Figure 45: Incremental budget impact – Scenario A versus Scenario B

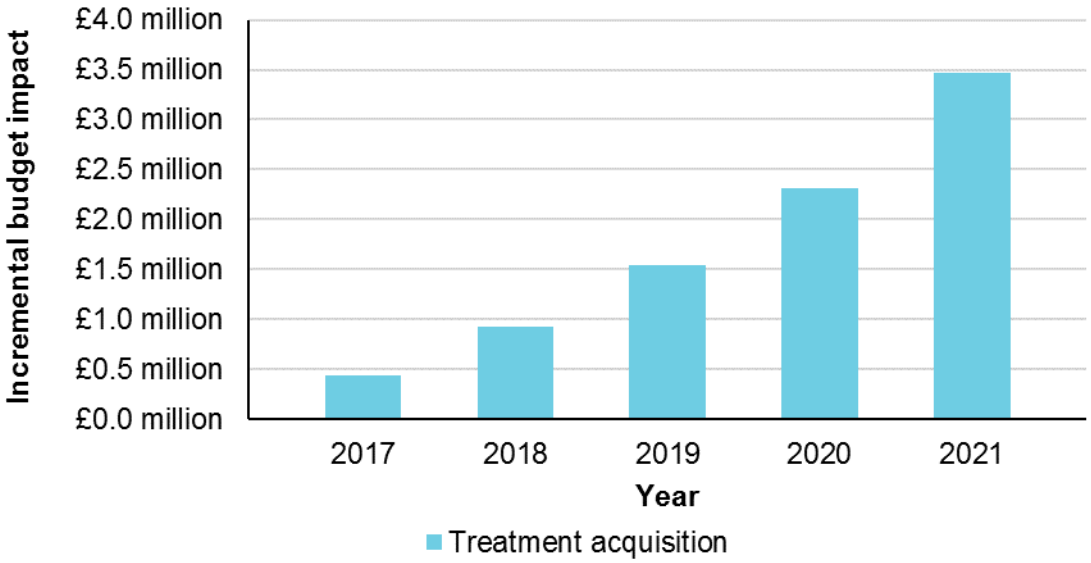
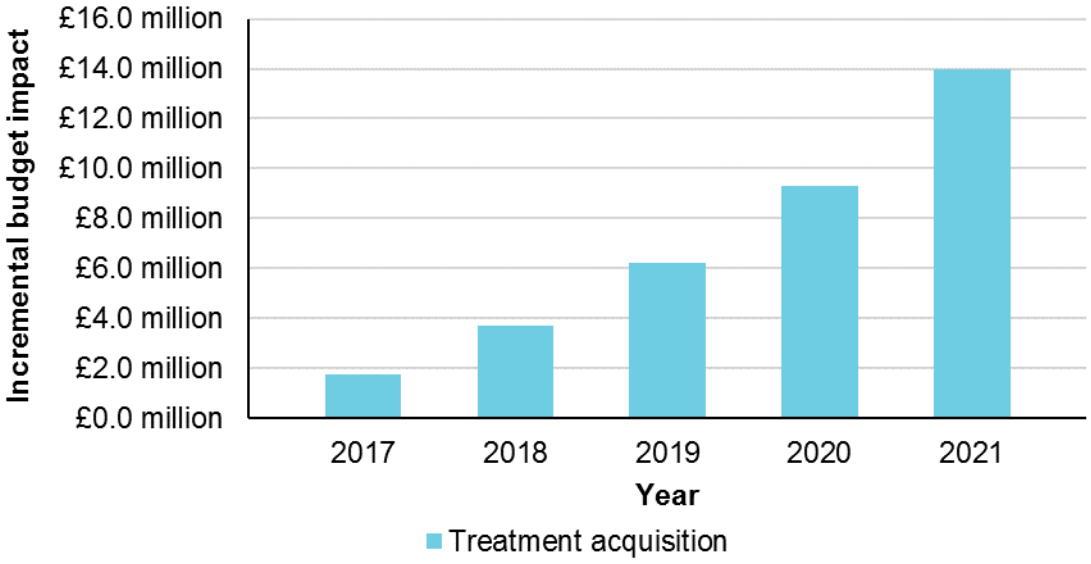


Figure 46: Incremental budget impact – Scenario A versus Scenario C



For Scenario A versus Scenario B, Figure 45 shows that NB32 is associated with an incremental budgetary impact of approximately £0.4 million in the first year it is made available, increasing to approximately £3.5 million by its fifth year of availability.

For Scenario A versus Scenario C, Figure 46 shows that NB32 is associated with an incremental budgetary impact of approximately £1.7 million in the first year it is made available, increasing to approximately £14.0 million by its fifth year of availability.

The increased budget impact should be considered in respect to the relatively large patient population and acknowledging that future obesity-related costs avoided are not captured within the analysis (given the restricted 5-year time horizon).

7 References

1. Greenway FL, Dunayevich E, Tollefson G, et al. Comparison of combined bupropion and naltrexone therapy for obesity with monotherapy and placebo. *J Clin Endocrinol Metab*. 2009; 94(12):4898-906.

2. Butland B, Jebb S, Kopelman P, et al. Tackling obesities: future choices - project report. 2007. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/287937/07-1184x-tackling-obesities-future-choices-report.pdf. Accessed: 11 October 2016.

3. 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.

4. World Health Organisation (WHO). Obesity and overweight. 2016. Available at: http://www.who.int/mediacentre/factsheets/fs311/en/. Accessed: 10 October 2016.

5. Haslam D. Obesity: current and emerging drug treatments. 2015. Available at: http://www.pharmaceutical-journal.com/learning/learning-article/obesity-current-and-emerging-drug-treatments/20068647.article. Accessed: 10 October 2016.

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

7. Gater DR, Jr. Obesity after spinal cord injury. *Phys Med Rehabil Clin N Am*. 2007; 18(2):333-51, vii.

8. Public Health England. Obesity and disability: Adults. 2013. Available at: http://www.noo.org.uk/uploads/doc/vid_18474_obesity_dis.pdf. Accessed: 13 October 2016.

9. Mills H. Psychological Consequences Of Being Overweight. 2005. Available at: https://www.mentalhelp.net/articles/psychological-consequences-of-being-overweight/. Accessed: 31 October 2016.

10. Roche. Xenical 120 mg hard capsules. 2009. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000154/WC500058428.pdf. Accessed: 26 July 2016.

11. Wharton S. Current Perspectives on Long-term Obesity Pharmacotherapy. *Can J Diabetes*. 2016; 40(2):184-91.

12. Genentech. Xenical prescribing information. 2015. Available at: https://www.gene.com/download/pdf/xenical_prescribing.pdf. Accessed: 16 November 2016.

13. Kahn SE, Zinman B, Haffner SM, et al. Obesity is a major determinant of the association of C-reactive protein levels and the metabolic syndrome in type 2 diabetes. *Diabetes*. 2006; 55(8):2357-64.

14. 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.

15. National Institute of Health and Care Excellence (NICE). Naltrexone-bupropion (prolonged release) for managing overweight and obesity: Final Scope. 2016. Available at: https://www.nice.org.uk/guidance/GID-TAG486/documents/final-scope. Accessed: 25 November 2016.

16. Greenway FL, Fujioka K, Plodkowski RA, et al. Effect of naltrexone plus bupropion on weight loss in overweight and obese adults (COR-I): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*. 2010; 376(9741):595-605.

17. Apovian CM, Aronne L, Rubino D, et al. A randomized, phase 3 trial of naltrexone SR/bupropion SR on weight and obesity-related risk factors (COR-II). *Obesity (Silver Spring)*. 2013; 21(5):935-43.

18. Wadden TA, Foreyt JP, Foster GD, et al. Weight loss with naltrexone SR/bupropion SR combination therapy as an adjunct to behavior modification: the COR-BMOD trial. *Obesity (Silver Spring)*. 2011; 19(1):110-20.

19. Hollander P, Gupta AK, Plodkowski R, et al. Effects of naltrexone sustained-release/bupropion sustained-release combination therapy on body weight and glycemic parameters in overweight and obese patients with type 2 diabetes. *Diabetes Care*. 2013; 36(12):4022-9.

20. Fujioka K, Plodkowski R, O'Neil PM, et al. The relationship between early weight loss and weight loss at 1 year with naltrexone ER/bupropion ER combination therapy. *Int J Obes (Lond)*. 2016; 40(9):1369-75.

21. Prieto L and Sacristán Jé A. What is the value of social values? The uselessness of assessing health-related quality of life through preference measures. *BMC Med Res Methodol*. 2004; 4:10.

22. Greenway FL, Whitehouse MJ, Guttadauria M, et al. Rational design of a combination medication for the treatment of obesity. *Obesity (Silver Spring)*. 2009; 17(1):30-9.

23. Vanina Y, Podolskaya A, Sedky K, et al. Body weight changes associated with psychopharmacology. *Psychiatr Serv*. 2002; 53(7):842-7.

24. Cowley MA, Pronchuk N, Fan W, et al. Integration of NPY, AGRP, and melanocortin signals in the hypothalamic paraventricular nucleus: evidence of a cellular basis for the adipostat. *Neuron*. 1999; 24(1):155-63.

25. Billes SK, Sinnayah P and Cowley MA. Naltrexone/bupropion for obesity: an investigational combination pharmacotherapy for weight loss. *Pharmacol Res*. 2014; 84:1-11.

26. Kelley AE and Berridge KC. The neuroscience of natural rewards: relevance to addictive drugs. *J Neurosci*. 2002; 22(9):3306-11.

27. Orexigen. Clinical Overview. 2015. Data on file.

28. European Medicines Agency (EMA). Assessment report for Xenical. 2012. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000154/WC500134995.pdf. Accessed: 26 October 2016.

29. Orexigen. Summary of Product Characteristics. Mysimba (Naltrexone/bupropion). 2016.

30. National Institute of Health and Care Excellence (NICE). Naltrexone-bupropion (prolonged release) for managing overweight and obesity: Response to consultee and commentator comments on the draft scope. 2016. Available at: https://www.nice.org.uk/guidance/GID-TAG486/documents/scope-consultation-comments-and-responses. Accessed: 8 November 2016.

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

32. James WP. The epidemiology of obesity: the size of the problem. *J Intern Med*. 2008; 263(4):336-52.

33. Baker C and Bate A. Obesity Statistics. 2016. Available at: researchbriefings.files.parliament.uk/documents/SN03336/SN03336.pdf. Accessed: 11 October 2016.

34. National Obesity Observatory. Obesity and Ethnicity. 2011. Available at: https://www.noo.org.uk/uploads/doc/vid_9444_Obesity_and_ethnicity_270111.pdf. Accessed: 31 October 2016.

35. Kearney J. Food consumption trends and drivers. *Philos Trans R Soc Lond B Biol Sci*. 2010; 365(1554):2793-807.

36. Health Communities. Risk Factors and Causes of Obesity. 2016. Available at: http://www.healthcommunities.com/obesity/causes.shtml. Accessed: 14 November 2016.

37. Diabetes UK. Number of people with diabetes reaches over 4 million. 2016. Available at: https://www.diabetes.org.uk/About_us/News/Number-of-people-with-diabetes-reaches-over-4-million/. Accessed: 31 October 2016.

38. Public Health England. Workplaces. 2016. Available at: https://www.noo.org.uk/LA/tackling/workplaces. Accessed: 13 October 2016.

39. Katz DA, McHorney CA and Atkinson RL. Impact of obesity on health-related quality of life in patients with chronic illness. *J Gen Intern Med*. 2000; 15(11):789-96.

40. Sullivan M, Karlsson J, Sjostrom L, et al. Swedish obese subjects (SOS)--an intervention study of obesity. Baseline evaluation of health and psychosocial functioning in the first 1743 subjects examined. *Int J Obes Relat Metab Disord*. 1993; 17(9):503-12.

41. Public Health England. Adult obesity and type 2 diabetes. 2014. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/338934/Adult_obesity_and_type_2_diabetes_.pdf. Accessed: 31 October 2016.

42. Public Health England. Social care. 2016. Available at: http://www.noo.org.uk/LA/impact/social. Accessed: 31 October 2016.

43. National Institute of Health and Care Excellence (NICE). PH53: Weight management: lifestyle services for overweight or obese adults. 2014. Available at: https://www.nice.org.uk/guidance/ph53. Accessed: 12 October 2016.

44. Health and Social Care Information Centre (HSIC). Health Survey for England Population Estimates Tables 2014. 2015. Available at: http://content.digital.nhs.uk/catalogue/PUB19297. Accessed: 21 November 2016.

45. European Association for the Study of Obesity (EASO). Obesity Facts & Figures. 2016. Available at: http://easo.org/education-portal/obesity-facts-figures/. Accessed: 13 October 2016.

46. House of Commons Health Committee. Obesity 2003-04. Available at: http://www.publications.parliament.uk/pa/cm200304/cmselect/cmhealth/23/23.pdf. Accessed: 13 October 2016.

47. National Institute of Health and Care Excellence (NICE). CG43: Obesity prevention. 2006. Available at: https://www.nice.org.uk/guidance/cg43. Accessed: 11 October 2016.

48. National Institute of Health and Care Excellence (NICE). QS127: Obesity: clinical assessment and management. 2016. Available at: https://www.nice.org.uk/guidance/qs127/chapter/List-of-quality-statements. Accessed: 12 October 2016.

49. National Institute of Health and Care Excellence (NICE). QS111: Obesity in adults: prevention and lifestyle weight management programmes. 2016. Available at: https://www.nice.org.uk/guidance/qs111. Accessed: 12 October 2016.

50. Rucker D, Padwal R, Li SK, et al. Long term pharmacotherapy for obesity and overweight: updated meta-analysis. *BMJ*. 2007; 335(7631):1194.

51. 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.

52. Nissen SE, Wolski KE, Prcela L, et al. Effect of Naltrexone-Bupropion on Major Adverse Cardiovascular Events in Overweight and Obese Patients With Cardiovascular Risk Factors: A Randomized Clinical Trial. *JAMA*. 2016; 315(10):990-1004.

53. European Medicines Agency (EMA). Assessment report for Mysimba (naltrexone/bupropion). 2014. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/003687/WC500185582.pdf. Accessed: 20 September 2016.

54. Halseth A, Shan K, Walsh B, et al. Method-of-Use Study of Naltrexone Sustained Release (SR)/Bupropion SR on Body Weight in Individuals with Obesity 17 November 2016 2016 (In Press). Data on file.

55. Committee for Medicinal Products for Human Use (CHMP). GUIDELINE ON CLINICAL EVALUATION OF MEDICINAL PRODUCTS USED IN WEIGHT CONTROL (CPMP/EWP/281/96 Rev. 1) 2007. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003212.pdf. Accessed: 20 September 2016.

56. Orexigen. A Multicenter, Randomized, Double Blind, Placebo Controlled Study Comparing the Safety and Efficacy of Two Doses of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Obese Subjects. (Clnical Study Report) 28 January 2010.

57. Rye K-A and Barter PJ. Cardioprotective functions of HDLs. *J Lipid Res*. 2014; 55(2):168-79.

58. Ridker PM, Danielson E, Fonseca FAH, et al. Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein. *N Engl J Med*. 2008; 359(21):2195-207.

59. Song Y, Manson JE, Tinker L, et al. Insulin Sensitivity and Insulin Secretion Determined by Homeostasis Model Assessment (HOMA) and Risk of Diabetes in a Multiethnic Cohort of Women: The Women's Health Initiative Observational Study. *Diabetes Care*. 2007; 30(7):1747-52.

60. Lee S-H, Kwon H-S, Park Y-M, et al. Predicting the Development of Diabetes Using the Product of Triglycerides and Glucose: The Chungju Metabolic Disease Cohort (CMC) Study. *PLoS ONE*. 2014; 9(2):e90430.

61. Public Health England. Measurement of Obesity. 2016. Available at: http://www.noo.org.uk/NOO_about_obesity/measurement. Accessed: 13 October 2016.

62. Orexigen. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Obese Subjects. (Clnical Study Report) 2010.

63. Orexigen. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Subjects with Obesity Participating in a Behavior Modification Program. (Clinical Study Report) 4 January 2010.

64. Orexigen. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone 32 mg Sustained Release/Bupropion 360 mg Sustained Release and Placebo in Obese Subjects with Type 2 Diabetes Mellitus. (Clinical Study Report) 2009.

65. R Core Team. The R Project for Statistical Computing. 2016. Available at: https://www.R-project.org/. Accessed: 22 November 2016.

66. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010; 36(3):1-48.

67. DerSimonian R and Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986; 7(3):177-88.

68. Higgins JP, Thompson SG, Deeks JJ and Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003; 327(7414):557-60.

69. 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.

70. Bakris G, Calhoun D, Egan B, et al. Orlistat improves blood pressure control in obese subjects with treated but inadequately controlled hypertension. *J Hypertens*. 2002; 20(11):2257-67.

71. 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.

72. Broom I, Wilding J, Stott P and Myers N. Randomised trial of the effect of orlistat on body weight and cardiovascular disease risk profile in obese patiensts: UK multimorbidity study. *Int J Clin Pract*. 2002; 56(7):494-9.

73. Derosa G, Mugellini A, Ciccarelli L and Fogari R. Randomized, double-blind, placebo-controlled comparison of the action of orlistat, fluvastatin, or both on anthropometric measurements, blood pressure, and lipid profile in obese patients with hypercholesterolemia prescribed a standardized diet. *Clin Ther*. 2003; 25(4):1107-22.

74. 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.

75. Gotfredsen A, Hendel HW and Andersen T. Influence of orlistat on bone turnover and body composition. *IntJ Obes Relat Metab Disord*. 2001; 25(8):1154-60.

76. 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.

77. 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.

78. Lindgarde F. The effect of orlistat body weight and coronary heart disease risk profile in obese patients: The swedish multimorbidity study. *J Intern Med*. 2000; 248(3):245-54.

79. Lucas CP, Boldrin MN and Reaven GM. Effect of orlistat added to diet (30% of calories from fat) on plasma lipids, glucose, and insulin in obese patients with hypercholesterolemia. *Am J Cardiol*. 2003; 91(8):961-4.

80. 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.

81. 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.

82. Swinburn BA, Carey D, Hills AP, et al. Effect of orlistat on cardiovascular disease risk in obese adults. *Diabetes Obes Metab*. 2005; 7(3):254-62.

83. Torgerson JS, Hauptman J, Boldrin MN and Sjöström L. XENical in the Prevention of Diabetes in Obese Subjects (XENDOS) Study. *Diabetes Care*. 2003; 27(1):155.

84. Joint Formulary Committee. British National Formulary (online) London: BMJ Group and Pharmaceutical Press. 2016. Available at: https://www.evidence.nhs.uk/formulary/bnf/current/4-central-nervous-system/45-obesity. Accessed: 14/09/16.

85. UK Prospective Diabetes Study Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of

complications in patients with type 2 diabetes (UKPDS 33). *Lancet*. 1998; 352(9131):837-53.

86. Lunn DJ, Thomas A, Best N and Spiegelhalter D. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput*. 2000; 10(4):325-37.

87. Brooks SP and Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat*. 1998; 7(4):434-55.

88. Woodward P. *Bayesian analysis made simple: An excel GUI for WinBUGS*. CRC Press, 2011.

89. Dias S, Welton NJ, Sutton AJ and Ades A. NICE DSU Technical Support Document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials. 2011. (Updated: September 2016) Available at: http://www.nicedsu.org.uk/TSD2%20General%20meta%20analysis%20corrected%202Sep2016v2.pdf. Accessed: 30/9/16.

90. Spiegelhalter DJ, Abrams KR and Myles JP. *Bayesian approaches to clinical trials and health-care evaluation*. John Wiley & Sons, 2004.

91. European Medicines Agency (EMA). Withdrawal of the marketing authorisation in the European Union. 2009. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Public_statement/2009/11/WC500012189.pdf. Accessed: 16 December 2016.

92. Orexigen. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Assessing the Occurrence of Major Adverse Cardiovascular Events (MACE) in Overweight and Obese Subjects With Cardiovascular Risk Factors Receiving Naltrexone SR/Bupropion SR. 11 May 2015. Data on file.

93. Khera R, Murad MH, Chandar AK, et al. Association of Pharmacological Treatments for Obesity With Weight Loss and Adverse Events: A Systematic Review and Meta-analysis. *JAMA*. 2016; 315(22):2424-34.

94. Watkins JB, Minshall ME and Sullivan SD. Application of economic analyses in U.S. managed care formulary decisions: A private payer's experience. *J Manag Care Pharm*. 2006; 12(9):726-35.

95. Veerman JL, Barendregt JJ, Forster M and Vos T. Cost-effectiveness of pharmacotherapy to reduce obesity. *PLoS ONE*. 2011; 6(10).

96. Salton R. Cost-effectiveness of bariatric surgery for the treatment of obesity in Australia. ISPOR 18th Annual European Congress. Milan, Italy. November 2015 2015. PSY73.

97. Ramirez M, Soto-Molina H, Rizzoli-Cordoba A, et al. Cost-effectiveness analisys of treatment with amfepramone (diethylpropion) in obesity in Mexico. *Value Health*. 2011; 14(3):A62-A3.

98. Molina HS, Martinez JPD, Juarez YE and Valle CFd. Cost-effectiveness analysis of amfepramone (diethylpropion) for the obesity treatment in Mexico. *ISPOR 17th Annual European Congress*. 2014.

99. Kwon JW, Song HJ, Lee HJ, et al. Cost-utility analysis of bariatric surgery for the treatment of obese patients in South Korea. *Value Health*. 2012; 15(4):A78.

100. Kim D, Kiebanoff MJ, Chhatwal J, et al. The cost-effectiveness of bariatric treatments in overweight and obese patients. ISPOR 21st Annual International Meeting. Washington D.C., U.S.A. May 2016 2016. PSY38.

101. Kalashnikova MF, Uchamprina VA, Romantsova TI and Gerasimov AN. Clinical and economic analysis of the modern strategies for treating metabolic syndrome. *Diabetes Mellitus*. 2014; 17(2):116-25.

102. Iannazzo S, Zaniolo O and Pradelli L. Economic evaluation of treatment with orlistat in Italian obese patients. *Curr Med Res Opin*. 2008; 24(1):63-74.

103. Hampp C, Hartzema AG and Kauf TL. Cost-utility analysis of rimonabant in the treatment of obesity. *Value Health*. 2008; 11(3):389-99.

104. Finkelstein EA, Kruger E and Karnawat S. Cost-effectiveness analysis of qsymia for weight loss (Provisional abstract). *PharmacoEconomics*. 2014; 33(7):699-706.

105. Finkelstein EA and Kruger E. Meta- and cost-effectiveness analysis of commercial weight loss strategies. *Obesity (Silver Spring)*. 2014; 22(9):1942-51.

106. Escudero GS, Idrovo J and Zapata L. Economic evaluation of metformin, metformin sibutramine or acarbose in the management of overweight and obese diabetes patients. *Value Health*. 2009; 12(3):A101.

107. Davies MJ, Chubb BD, Smith IC and Valentine WJ. Cost-utility analysis of liraglutide compared with sulphonylurea or sitagliptin, all as add-on to metformin monotherapy in Type 2 diabetes mellitus (Structured abstract). *Diabet Med*. 2012; 29.

108. Burch J, McKenna C, Palmer S, et al. Rimonabant for the treatment of overweight and obese people. *Health Technol Assess*. 2009; 13 Suppl 3:13-22.

109. Beaudet A, Palmer JL, Timlin L, et al. Cost-utility of exenatide once weekly compared with insulin glargine in patients with type 2 diabetes in the UK. *J Med Econ*. 2011; 14(3):357-66.

110. Baal PH, Berg M, Hoogenveen RT, et al. Cost-effectiveness of a low-calorie diet and orlistat for obese persons: modeling long-term health gains through prevention of obesity-related chronic diseases (Structured abstract). *Value Health*. 2008; 11(Issue):1033-40.

111. Ara R and Brennan A. The cost-effectiveness of sibutramine in non-diabetic obese patients: evidence from four Western countries. *Obes Rev*. 2007; 8(4):363-71.

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

113. Medicines & Healthcare products Regulatory Agency. The Clinical Practice Research Datalink. 2016. Available at: https://www.cprd.com/home/. Accessed: 10 November 2016.

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

115. National Institute of Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. 2013. Available at: https://www.nice.org.uk/process/pmg9/chapter/foreword. Accessed: 1 September 2016.

116. Levy DT, Mabry PL, Wang YC, et al. Simulation Models of Obesity: A Review of the Literature and Implications for Research and Policy. *Obes Rev*. 2011; 12(5):378-94.

117. Dare S, Mackay DF and Pell JP. Relationship between Smoking and Obesity: A Cross-Sectional Study of 499,504 Middle-Aged Adults in the UK General Population. *PLoS ONE*. 2015; 10(4):e0123579.

118. Office for National Statistics. National Life Tables: United Kingdom. 2016. (Updated: 29 September 2016) Available at: http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables. Accessed: 22 November 2016.

119. Kolotkin RL, Meter K and Williams GR. Quality of life and obesity. *Obes Rev*. 2001; 2.

120. Kolotkin RL, Crosby RD, Williams GR, et al. The relationship between health-related quality of life and weight loss. *Obes Res*. 2001; 9(9):564-71.

121. Anandacoomarasamy A, Caterson ID, Leibman S, et al. Influence of BMI on health-related quality of life: Comparison between an obese adult cohort and age-matched population norms. *Obesity (Silver Spring)*. 2009; 17(11):2114-8.

122. Baxter S, Sanderson K, Venn A, et al. Construct validity of SF-6D health state utility values in an employed population. *Qual Life Res*. 2015; 24(4):851-70.

123. Boesten JEJ, Kaper J, Stoffers HEJH, et al. Rimonabant improves obesity but not the overall cardiovascular risk and quality of life; results from CARDIO-REDUSE (CArdiometabolic Risk reDuctiOn by Rimonabant: The effectiveness in daily practice and its USE). *Family Practice*. 2012; 29(5):521-7.

124. Boye KS, Yurgin N, Dilla T, et al. Health-related quality of life of patients with type 2 diabetes mellitus in primary care in Spain: Self-reported and proxy assessment using the EQ-5D. *J Med Econ*. 2007; 10(1):41-58.

125. Clemens S, Begum N, Harper C, et al. A comparison of EQ-5D-3L population norms in Queensland, Australia, estimated using utility value sets from Australia, the UK and USA. *Qual Life Res*. 2014; 23(8):2375-81.

126. 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).

127. Doyle S, Lloyd A, Birt J, et al. Willingness to pay for obesity pharmacotherapy. *Obesity (Silver Spring)*. 2012; 20(10):2019-26.

128. 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.

129. Goren A and Philip EJ. Obesity associated with added burden among patients with cancer in the 2012 national health and wellness survey in the united states. *Value Health*. 2013; 16(3):A145.

130. Gupta S, Wang Z and Flores NM. The incremental burden of obesity and sleep disorders among individuals with type 2 diabetes mellitus. ISPOR 20th Annual International Meeting. Philadelphia, PA, U.S.A. May 2015 2015. PDB83.

131. Hunger M, Schunk M, Meisinger C, et al. Estimation of the relationship between body mass index and EQ-5D health utilities in individuals with type 2 diabetes: Evidence from the population-based KORA studies. *J Diabetes Complications*. 2012; 26(5):413-8.

132. Jia H, Zack MM, Moriarty DG and Fryback DG. Predicting the EuroQol Group's EQ-5D index from CDC's healthy days in a US sample. *Med Decis Making*. 2011; 31(1):174-85.

133. Keating CL, Peeters A, Swinburn BA, et al. Utility-based quality of life associated with overweight and obesity: the Australian diabetes, obesity, and lifestyle study. *Obesity (Silver Spring)*. 2013; 21(3):652-5.

134. 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.

135. Kolotkin RL, Smolarz G, Mein HH and Bjorner JB. Improvement in IWQOL-Lite-derived health utility score with liraglutide 3.0 mg versus placebo over 3 years in prediabetes. *ISPOR 21st Annual International Meeting*. 2016.

136. Koo EB, Enamandram M and Kimball AB. Relative to obesity, negative effect of hidradenitis suppurativa is enhanced on quality of life in super obese patients. *J Invest Dermatol*. 2014; 134:S104.

137. Korhonen PE, Seppälä T, Järvenpää S and Kautiainen H. Body mass index and health-related quality of life in apparently healthy individuals. *Qual Life Res*. 2014; 23(1):67-74.

138. Lacey LA, Wolf A, O'Shea D, et al. Cost-effectiveness of orlistat for the treatment of overweight and obese patients in Ireland. *Int J Obes (Lond)*. 2005; 29(8):975-82.

139. Maheswaran H, Petrou S, Rees K and Stranges S. Estimating EQ-5D utility values for major health behavioural risk factors in England. *J Epidemiol Community Health*. 2013; 67(2):172-80.

140. McDonough C, Dunkley AJ, Aujla N, et al. The association between body mass index and health-related quality of life: Influence of ethnicity on this relationship. *Diabetes Obes Metab*. 2013; 15(4):342-8.

141. Norman R, Church J, Berg Bvd and Goodall S. Australian health-related quality of life population norms derived from the SF-6D. *Aust N Z J Public Health*. 2013; 37(1):17-23.

142. Oh SH, Song HJ, Lee HJ, et al. The improvement of quality of life in patients treated with bariatric surgery. *Value Health*. 2012; 15(4):A79.

143. Orlic Z, Dibonaventura MD, Gupta S, et al. Health outcomes among type 2 diabetes patients with comorbid obesity in Brazil. *Value Health*. 2012; 15(4):A184.

144. Pinto AM, Subak LL, Nakagawa S, et al. The effect of weight loss on changes in health-related quality of life among overweight and obese women with urinary incontinence. *Qual Life Res*. 2012; 21(10):1685-94.

145. Richard L, Gupta S, Pomerantz D and Forsythe A. Is obesity a problem in Brazil? . ISPOR 17th Annual European Congress. Amsterdam, The Netherlands. November 2014 2014. PSY48.

146. Rothberg AE, McEwen LN, Kraftson AT, et al. The impact of weight loss on health-related quality-of-life: implications for cost-effectiveness analyses. *Qual Life Res*. 2014; 23(4):1371-6.

147. Rubin RR, Peyrot M, Wang NY, et al. Patient-reported outcomes in the practice-based opportunities for weight reduction (POWER) trial. *Qual Life Res*. 2013; 22(9):2389-98.

148. Sartorio A, Agosti F, Col AD, et al. Concurrent comparison of the measurement properties of generic and disease-specific questionnaires in obese inpatients. *J Endocrinol Invest*. 2014; 37(1):31-42.

149. Serrano-Aguilar P, Muñoz-Navarro SR, Ramallo-Fariña Y and Trujillo-Martín MM. Obesity and health related quality of life in the general adult population of the Canary Islands. *Qual Life Res*. 2009; 18(2):171-7.

150. 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.

151. Tsai AG, Wadden TA, Volger S, et al. Cost-effectiveness of a primary care intervention to treat obesity. *Int J Obes (Lond)*. 2013; 37(SUPPL.1):S31-S7.

152. Vetter ML, Wadden TA, Lavenberg J, et al. Relation of health-related quality of life to metabolic syndrome, obesity, depression and comorbid illnesses. *Int J Obes (Lond)*. 2011; 35(8):1087-94.

153. Wang L, Wu YQ, Tang X, et al. Profile and correlates of health-related quality of life in Chinese patients with coronary heart disease. *Chin Med J (Engl)*. 2015; 128(14):1853-61.

154. Wee CC, Davis RB, Chiodi S, et al. Sex, race, and the adverse effects of social stigma vs. other quality of life factors among primary care patients with moderate to severe obesity. *J Gen Intern Med*. 2014; 30(2):229-35.

155. Wee CC, Davis RB, Huskey KW, et al. Quality of life among obese patients seeking weight loss surgery: The importance of obesity-related social stigma and functional status. *J Gen Intern Med*. 2013; 28(2):231-8.

156. Wiklund M and Olsén MF. Physical function and health related quality of life before and 18 months after bariatric surgery. *Open Obes J*. 2015; 7(1):42705.

157. Yoon YS, Lee ES, Lee JH and Yang YJ. Obesity related quality of life in Korea: Data from the 2005 Korea National Health and Nutrition Examination Survey (KNHANES). *Obes Rev*. 2010; 11:316.

158. Zhang P, Hire D, Espeland MA, et al. Impact of intensive lifestyle intervention on preference-based quality of life in type 2 diabetes: Results from the Look AHEAD trial. *Obesity (Silver Spring)*. 2016; 24(4):856-64.

159. Hernández Alava M, Wailoo A and Ara R. Tails from the Peak District: adjusted censored mixture models of EQ-5D health state utility values. *School of Health and Related Research, University of Sheffield, Health Economics & Decision Science Discussion Papers*. 2010; HEDS Discussion Paper 10/08.

160. 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.

161. Wahba J, Jasmat I, Patel A, et al. An audit into the management of obesity in pregnancy at a major London teaching hospital. *BJOG*. 2015; 122:229.

162. Srishanmuganathan J, Patel H, Car J and Majeed A. National trends in the use and costs of anti-obesity medications in England 1998-2005. *J Public Health (Oxf)*. 2007; 29(2):199-202.

163. Scarborough P, Bhatnagar P, Wickramasinghe KK, et al. The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: An update to 2006-07 NHS costs. *J Public Health (Oxf)*. 2011; 33(4):527-35.

164. Reeves GK, Balkwill A, Cairns BJ, et al. Hospital admissions in relation to body mass index in UK women: A prospective cohort study. *BMC Medicine*. 2014; 12(1).

165. Pollock RF, Chilcott J, Muduma G and Valentine WJ. Laparoscopic adjustable gastric banding vs standard medical management in obese patients with type 2 diabetes: a budget impact analysis in the UK. *J Med Econ*. 2013; 16(2):249-59.

166. Murray HL, Dawes T, Nicholas CR and Margarson MP. Effect of an enhanced recovery programme on length of stay in gastric bypass surgery. *Eur J Anaesthesiol*. 2011; 28:22-3.

167. Meads DM, Hulme CT, Hall P and Hill AJ. The cost-effectiveness of primary care referral to a UK commercial weight loss programme. *Clinical Obesity*. 2014; 4(6):324-32.

168. McEwan P, Gordon J, Evans M and Puelles J. Factors associated with weight gain and hypoglycaemia and the impact upon hospitalisation in type 2 diabetes patients managed with metformin plus sulphonylurea. ISPOR 17th Annual European Congress. Amsterdam, The Netherlands. November 2014 2014. PDB161.

169. Lewis L, Taylor M, Broom J and Johnston KL. The cost-effectiveness of the LighterLife weight management programme as an intervention for obesity in England. *Clinical Obesity*. 2014; 4(3):180-8.

170. Lean M, Brosnahan N, McLoone P, et al. Feasibility and indicative results from a 12-month low-energy liquid diet treatment and maintenance programme for severe obesity. *Br J Gen Pract*. 2013; 63(607):e115-e24.

171. Flanagan C, Ryan M and Reekie S. Motivate: A structured behavioural weight management program for obese adults with long term condition (LTC) is a cost-effective treatment. *Obes Facts*. 2013; 6:112.

172. Diaz-Cuervo H, Callejo D, Cuervo J, et al. Costs of management of pregnancy in obese women across Europe: The DALI experience. *Value Health*. 2013; 16(7):A636.

173. 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.

174. CounterweightProjectTeam. Influence of body mass index on prescribing costs and potential cost savings of a weight management programme in primary care. *J Health Serv Res Policy*. 2008; 13(3):158-66.

175. Capehorn M, Carter D and Senior L. Service evaluation of the rotherham institute for obesity. *Obes Facts*. 2012; 5:187.

176. Bush J, Langley C, Mills S and Hindle L. A comparison of the provision of the My Choice Weight Management Programme via general practitioner practices and community pharmacies in the United Kingdom. *Clinical Obesity*. 2014; 4(2):91-100.

177. MIMS. Orlistat. 2016. Available at: http://www.mims.co.uk/drugs/nutrition/obesity/orlistat. Accessed: 21 Nov 2016.

178. Health and Social Care Information Centre (HSIC). Prescription Cost Analysis: England 2015 - Data by individual preparation - Orlistat. 2015.

179. Curtis L and Burns A. Unit Costs of Health and Social Care. University of Kent: Personal Social Services Research Unit, 2015.

180. 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.

181. Ara R, Pandor A, Stevens J, et al. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. *Health Technol Assess*. 2009; 13(34):1-74, 5-118.

182. Robinson M, Palmer S, Sculpher M, et al. Cost-effectiveness of alternative strategies for the initial medical management of non-ST elevation acute coronary syndrome: systematic review and decision-analytical modelling. *Health Technol Assess*. 2005; 9(27):iii-iv, ix-xi, 1-158.

183. Youman P, Wilson K, Harraf F and Kalra L. The Economic Burden of Stroke in the United Kingdon. *PharmacoEconomics*. 2003.

184. Clarke P, Gray A and Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). *Med Decis Making*. 2002; 22(4):340-9.

185. Gillett M, Dallosso HM, Dixon S, et al. Delivering the diabetes education and self management for ongoing and newly diagnosed (DESMOND) programme for people with newly diagnosed type 2 diabetes: cost effectiveness analysis. *BMJ*. 2010; 341:c4093.

186. Diabetes UK. Cost of Diabetes 2016. 2016. Available at: http://www.diabetes.co.uk/cost-of-diabetes.html. Accessed: 8 Nov 2016.

187. World Health Organisation (WHO). Cardiovascular diseases (CVDs). 2016. Available at: http://www.who.int/mediacentre/factsheets/fs317/en/. Accessed: 11 Oct 2016.

188. World Health Organisation (WHO). Types of Cardiovascular Disease. 2004. Available at: http://www.who.int/cardiovascular_diseases/en/cvd_atlas_01_types.pdf. Accessed: 22 November 2016.

189. Bhaskaran K, Douglas I, Forbes H, et al. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *Lancet*. 2014; 384(9945):755-65.

190. BNF Online. Adverse Events. 2016. Available at: http://www.evidence.nhs.uk/formulary/bnf/current/guidance-on-prescribing/adverse-reactions-to-drugs. Accessed: 22 Sept 2016.

191. Ul-Haq Z, Mackay DF, Fenwick E and Pell JP. Impact of metabolic comorbidity on the association between body mass index and health-related quality of life: a Scotland-wide cross-sectional study of 5,608 participants. *BMC Public Health*. 2012; 12(1):143.

192. Health and Social Care Information Centre (HSIC). HSIC Quality and Outcomes Framework (QOF) for April 2014 - March 2015, England. GPES and CQRS database - 2014-15 data extracted 10th July 2015. 2015. (Updated: 10 July 2015) Available at: http://content.digital.nhs.uk/catalogue/PUB18887. Accessed: 15 December 2016.

8 Appendices

The following appendices are included in a separate appendices document.

Appendix 1: European public assessment report, SmPC/IFU, scientific discussion or drafts (Section 2.2)

Appendix 2: Search strategy for relevant RCT studies (Section 4.1)

Appendix 3: Statistical analysis, participant flow and quality assessment of the NB-CVOT study

Appendix 4: Quality assessment of randomised controlled trials (RCTs) (Section 4.6)

Appendix 5: Sensitivity analyses and sub-studies

Appendix 6: Patient reported outcome and health-related quality of life tools

Appendix 7: Primary and secondary outcomes of the NB-CVOT study (Section 4.7)

Appendix 8: Frequentist pairwise meta-analysis (Section 4.9.3)

Appendix 9: SLR-identified studies excluded from ITC analyses (Section 4.10.2)

Appendix 10: Data imputation (Section 4.10.3)

Appendix 11: Trial baseline characteristics (Section 4.10.3)

Appendix 12: RCT quality assessment for the orlistat RCT's included NMA (Section 4.10.4)

Appendix 13: Frequentist pairwise meta-analysis (Section 4.10.5)

Appendix 14: WinBUGS Code (Section 4.10.5)

Appendix 15: Published cost-effectiveness studies (Section 5.1)

Appendix 16: Measurement and valuation of health effects (Section 5.4)

Appendix 17: Cost and healthcare resource use identification measurement and valuation studies (Section 5.5)

Appendix 18: Summary of base case *de novo* analysis inputs and assumptions

Appendix 19: Overview of the running of probabilistic sensitivity analysis in the *de novo* economic model

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

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

Dear Hans-Joerg,

The Evidence Review Group, Kleijnen systematic reviews, and the technical team at NICE have looked at the submission received on 5 January 2017 from Orexigen Therapeutics. In general they felt that it is well presented and clear. However, the ERG and the NICE technical team would like further clarification on the clinical and cost effectiveness data (see questions listed at end of letter).

The ERG and the technical team at NICE will be addressing these issues in their reports.

Please provide your written response to the clarification questions by **5pm** on **13 February 2017**. Your response and any supporting documents should be uploaded to NICE Docs/Appraisals https://appraisals.nice.org.uk/request/24165.

Two versions of your written response should be submitted; one with academic/commercial-in-confidence information clearly marked and one with this information removed.

Please underline all confidential information, and separately highlight information that is submitted as commercial in confidence in turquoise, and all information submitted as academic in confidence in yellow.

If you present data that are not already referenced in the main body of your submission and that are academic/commercial in confidence, please complete the attached checklist for confidential information.

Please do not embed documents (PDFs or spreadsheets) in your response because this may result in them being lost or unreadable.

If you have any queries on the technical issues raised in this letter, please contact Hamish Lunagaria, Technical Lead (Hamish.lunagaria@nice.org.uk). Any procedural questions should be addressed to Liv Gualda Project Manager (liv.gualda@nice.org.uk).

Yours sincerely

Joanna Richardson
Technical Adviser – Appraisals
Centre for Health Technology Evaluation

Encl. checklist for confidential information

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Section A: Clarification on effectiveness data

Literature searching

A1. Regarding the Medline/Embase strategies reported for all sections, please clarify if this was a single search conducted simultaneously over both the Embase and Medline individual databases or was it a single search of Embase conducted on the understanding that it now contains all records from Medline.

A2. Please provide search dates for the conference searches for all sections and provide details of any search strategies used.

A3. The ERG are concerned by the use of study design filters in the Cochrane Library searches of CDSR, DARE and CENTRAL listed in Appendix 2. We feel that this is an overly restrictive approach given that these resources are already filtered by study design. We reran your strategy and noted that the combination of Lines #11 and #17 when limited to CENTRAL brought back 338 records, however the additional study design filters which appeared in your strategy reduced this to between 268-273 depending on whether you use your limit code of "and CCRCT" (line #92) or Wiley's own Trials limit. Please rescreen these results missed by your study design filters and confirm whether any records meeting your inclusion criteria were missed by this approach.

A4. Section 4.12 talks about adverse events, however no mention is made of how this information was identified and no searches were reported. Please confirm which searches were used to inform this section. If the searches reported in Appendix 2 were used please confirm if all of the results retrieved were screened for adverse events. If additional searches were used, please provide full details.

Furthermore, please confirm that only RCTs were included, even for adverse events.

A5. Please confirm that the Econlit search was carried out on the EBSCO platform as stated in Table 22, Appendix 15. It is our understanding that EBSCO host does not support the search of MH as a field in Econlit as shown in your strategy.

A6. The Econlit search in Appendix 15 appears to contain an error in the line numbers being combined in lines S60 and S61. The line above (S59) has the combination "S11 AND S25 AND S58" which appears to be correct; however the following two lines have the combination "S11 AND S22 AND S58. Line S25 is a combination of all listed interventions where line S22 is for "TI (lorcaserin OR belviq) OR AB (lorcaserin OR belviq)". Please confirm whether this was an error in reporting or one which occurred during the search.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Outcomes

A7. Body Mass Index (BMI)

a. Please justify why BMI was not evaluated as an outcome in the submission according to the NICE scope?

b. **Priority**: Mean BMI at baseline is provided for each of the four main naltrexone-bupropion trials. Please calculate mean BMI at week 56 (or end-of-study) using the height at baseline for all four trials (COR-I, COR-II, COR-BMOD and COR-DM).

Clinical trial data and results

A8. In section 2.2 on page 26 of the company submission an ongoing phase IV study is mentioned on the occurrence of MACE in overweight and obese patients taking NB32. Please provide bibliographical details of the trial? Please provide the protocol. Are any interim data available? If so then please provide these?

A9. Please provide bibliographic details and the protocol of the ongoing trial evaluating safety of NB32 in patients with renal or hepatic impairment mentioned in section 2.2 on page 26?

A10. **Priority**: In section 2.3 table 6 there is a statement “Retreatment with NB32 is not routinely anticipated and thus not modelled.” Please justify why patients would not be retreated with naltrexone-bupropion for any subsequent weight gain after a successful treatment with the drug?

A11. **Priority**: In section 2.3 table 6 there is a statement ‘For patients continuing treatment post 16 weeks, treatment should be continued as long as clinical benefit is observed.’ The main trials are just over a year’s duration.

a. Please provide the precise criteria by which treatment discontinuation was determined in the trials?

b. Please clarify what was the percentage of participants in each of the trials that discontinued due to cessation of clinical benefit?

c. Please provide the precise criteria by which treatment discontinuation would be determined in clinical practice?

A12. **Priority**: Please justify not including standard management as an intervention in the review eligibility criteria and the searches? Currently standard care is considered only as a comparator to orlistat or naltrexone-bupropion. Therefore, any studies comparing behavioural interventions with no treatment (or other behavioural interventions) are excluded. However, these are relevant according to the scope.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

A13. Please clarify why non-RCTs were eligible for the review but not considered further, not even for adverse events?

A14. Table 10, Eligibility criteria for trials to be included in the systematic review, (CS, page 44) mentions that 'Studies published in non-English languages were flagged'. Please explain what is meant by that and please explain what was done with these studies.

A15. Please extend the flow chart in Figure 2 of the submission to illustrate the studies considered in the direct meta-analysis and those forming part of the indirect comparison with reasons for exclusion?

A16. Please clarify how many, if any, of the patients in the COR trials and NB-CVOT had previously received treatment with orlistat?

A17. Please provide a summary table of percentages of patients in the COR trials and NB-CVOT who are both overweight according to the NICE scope (≥ 27 kg/m^2 to < 30 kg/m^2) and with one or more weight-related co-morbidities?

A18. **Priority**: Please describe in more detail the components of standard care in the four COR trials? Please include summary statistics of number of contacts with each type of health care professional as well as any specific instructions to exercise or to attend a weight loss club. How was consistency of standard care between centres within a trial assured?

A19. ITT analysis

a. Please justify the use of a modified ITT analysis in the COR trials?

b. **Priority**: Please provide all clinical effectiveness outcomes from the NB trials used in the economic model based on two ITT populations: "ITT with the Weight regain imputation method" and "ITT with Baseline-carried forward analysis". In other words, please provide all data for NB32 and SM as reported in CS Tables 56-58 (proportion of responders, average weight loss at the assessment moments) for these two ITT populations. Please also provide data on treatment discontinuation (before, between and after the two assessments moments) based on these two populations.

c. **Priority**: Please provide all clinical effectiveness outcomes from the COR-BMOD trial for the control arm (intensive behaviour modification only) if used in the economic model based on two ITT populations: "ITT with the Weight regain imputation method" and "ITT with Baseline-carried forward analysis". In other words, please provide all data from the control arm of COR-BMOD as reported in CS Tables 56-58 (proportion of responders, average weight loss at the assessment moments) for these two ITT

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

populations. Please also provide data on treatment discontinuation (before, between and after the two assessments moments) based on these two populations.

A20. Please justify why only 5% reduction in weight and mean % weight change from baseline at 1 year were chosen as outcomes for the meta-analysis?

A21. Please provide four clinical effectiveness outcomes (Mean % weight change from baseline at 1 year; 5% reduction in weight at 1 year from baseline; Change in waist circumference (cm) at 1 year; and Proportion of patients with ≥10% decrease in body weight at 1 year (the 1-year time point ranging from 52 to 57 weeks)) from four NB32 trials (COR-I, COR-II, COR-BMOD and COR-DM) based on two ITT populations: "ITT with the Weight regain imputation method" and "ITT with Baseline-carried forward analysis". And please provide the same meta-analyses results based on these two ITT populations as reported in chapter 4.9 in the CS.

Section B: Clarification on cost-effectiveness data

Treatment effectiveness

B1. **Priority**: Ara et al.[1] state that 'although there was a wide variation in the modelling approaches and evidence used in the studies, the variable reported to have the largest effect on the results in the majority of the models was the period of weight regain modelled.'[1] Several assumptions for weight regain in the CS base-case were discrepant with assumptions from the base-case analysis by Ara et al.[1] (see CS Table 53)

a. Please justify why weight regain towards the predicted BMI (with the natural history model) was preferred over weight regain towards the baseline BMI.

b. Please justify why the period of three years (for linear weight regain) is appropriate.

c. Please justify why weight regain towards the predicted BMI (with the natural history model) was only started after discontinuation of **all treatments** instead of after discontinuation of **active treatments** as assumed by Ara et al.[1]

d. Please provide a scenario analysis assuming start of weight regain after discontinuation of **active treatments**.

e. Please provide a scenario analysis, similar to Ara et al.'s[1] base case, in which patients revert to their baseline BMI in three years and then enter the natural history model.

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

f. Please provide a scenario analysis combining d and e: patients revert to their baseline BMI in three years and then enter the natural history model and start of weight regain after discontinuation of active treatments.

B2. **Priority**: The 16 weeks treatment discontinuation for NB32 was linearly scaled to 12 weeks and assumed to be equivalent to treatment discontinuation used for orlistat.

a. Please justify why the treatment discontinuation for NB32 (linearly scaled or not linearly scaled) is applicable to orlistat.

b. Please provide a scenario analysis using the NB32 treatment discontinuation for orlistat without linear scaling.

c. At the end of the trial follow-up period, it is assumed that all patients would discontinue treatment. Please justify this assumption further and provide a scenario analysis using parametric survival models applied to the COR trial data to extrapolate treatment discontinuation.

d. Please clarify what determines whether patients can continue standard management after they have discontinued pharmacological treatment and clarify how time to discontinuation of standard management is subsequently estimated for these patients.

e. Please justify why it was appropriate to use the NB-CVOT study to estimate treatment discontinuation of standard management beyond 56 weeks despite the difference in population compared with the COR trial programme (which was used for estimating treatment discontinuation of standard management up to 52 weeks). Please also discuss the implications of using a more severe patient population for estimating treatment discontinuation post 56 weeks.

B3. **Priority**: No re-treatment or alternative treatments after treatment discontinuation are assumed in the model.

a. Please justify the assumption of no re-treatment after treatment discontinuation and provide a scenario analysis incorporating re-treatment with active treatments (i.e. NB32 and/or orlistat) and another scenario analysis incorporating re-treatment with standard management.

b. Please justify the assumption of no alternative treatments after treatment discontinuation and provide a scenario analysis incorporating alternative treatments (e.g. bariatric surgery).

B4. There is no justification in the company submission for why baseline patient characteristics of patients who receive aspirin and patients who receive anti-hypertensive medication are not varied in the generation of profiles in the CS. In the

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

model, the justification reads that these settings 'are disabled as the risk equations by Ara et al. (2012)[1] cause counter-intuitive results (for example, an increase in BMI causing a decrease in the time to death).' However, the ERG would like to highlight that it seems plausible that an increase in BMI would cause a decrease in the time to death.

a. Please provide clarification and justification for this?

b. Please provide a scenario analysis in which these parameters are allowed to vary?

B5. In the company submission it is stated that 'mean change in body weight estimates determines the proportion of responders and non-responders at secondary response.' However, after the primary assessment, responders and non-responders are assigned a mean change in body weight. Specifically, responders at the first assessment for NB32 are assigned an average weight loss of 9.4%. Hence, these responders at the first assessment automatically also meet the response criterion (i.e. ≥5% weight loss) for the second assessment. In other words, NB32 responders at the first assessment are also automatically responders on the second assessment, if they continue treatment.

a. Please clarify how the proportion of responders and non-responders at the secondary assessment are incorporated in the model.

b. Please clarify whether for NB32 and orlistat, responders at the first assessment are also automatically responders at the second assessment if they continue treatment. If this is the case, justify this assumption and provide a scenario analysis allowing patients to be identified as non-responders at the second assessment.

B6. **Priority**: Please provide two scenario analyses using data on clinical effectiveness and treatment discontinuation derived from the two ITT populations described in Question A13b from the Clinical Effectiveness section: one based on the ITT with weight regain imputation method; and one based on the ITT with baseline-carried forward analysis.

B7. In the company submission model, at diabetic onset (and stroke / MI events), time to primary and secondary assessment is recalculated without subtracting the time at which this event occurred. This appears to delay the time to assessment for those patients that experienced the onset of the respective event before either one of the assessments.

a. Please justify why time to assessment was recalculated in this way?

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

b. If the time to assessment was, in fact, a mistake, please provide results of a corrected analysis?

Comparators

B8. **Priority**: Please add intense behavioural modification as a comparator in the model and provide cost-effectiveness results? Please use the responses to clarification question A13c as well as modified resource use and costs data to reflect intense behavioural modification as a comparator in the model.

Model structure

B9. It is assumed that only 2 strokes, 2 MIs or 1 stroke and 1 MI can occur (with or without T2DM and patients can develop T2DM after the first event). Please justify that this simplifying assumption is plausible, e.g. that a stroke after 2 MIs does not have any important costs and-or quality of life implications?

B10. General population mortality data are used to inform the probability of death beyond follow-up of 15 years. Please justify this assumption and provide an alternative scenario analysis without this assumption.

Health related quality of life

B11. Utility scores are derived from a Tobit model from PHE.

a. Please clarify that the utility scores obtained from this model have face validity, e.g. by means of provision and discussion of a table with utility scores associated with experiencing the different (combinations of) health events in the model, for an average patient.

b. Please justify that the Tobit model was preferred over the OLS regression model (CS Table 60).

B12. Please provide a scenario analysis using the SF-36 data from the COR-II trial.

B13. Please provide justification for why no utility decrements were applied to adverse events.

Resource use and costs

B14. In the company submission, the cost of Diabetes Mellitus are not derived from Ara et al.[1] Please justify this.

B15. Please justify why drug wastage of NB32 was not incorporated in the model.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

Probabilistic sensitivity analysis

B16. Please provide a justification as to why the company believes that probabilistic sensitivity analysis using 100 simulations results in stable / plausible results.

Cost effectiveness results

B17. Please provide an overview of the disaggregated costs, QALYs and LYs (using the conditions specified in company submission Figure 25).

B18. Ara et al.[1] used a cohort of 1,000,000 patients in their patient-level simulation and stated that, with a cohort size of 200,000 patients, there was still a small amount of variation in results, which stabilised after simulation of 400,000 patients. In contrast, a cohort of only 1,000 patients was used in the company submission. Company submission Figures 34 and 35 provide a diagnostic exercise to examine the minimum number of patients needed to obtain stable results.

a. Please provide similar figures using the incremental costs, incremental QALYs and the ICER (QALYs) and justify why 1,000 patients were deemed sufficient.

b. Please justify the usage of 1,000 patients given that Ara et al[1] used a cohort of 1,000,000 patients and stated that, with a cohort size of 200,000 patients, there was still a small amount of variation in results.

Validity

B19. Please provide the results of the internal validation described at the end of company submission section 5.10.

B20. Please provide the source for and justify the validity of equation 1 in the company submission. Additionally, provide a simple example using this formula and explain why the results are plausible.

Section C: Textual clarifications and additional points

None

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.

NICE National Institute for Health and Care Excellence

10 Spring Gardens
London
SW1A 2BU
United Kingdom

+44 (0)300 323 0140

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

Dear Liv,

Please find enclosed Orexigen Therapeutics response to the clarification questions from the Evidence Review Group, Kleijnen systematic reviews, received on the 30 January 2017.

Please let me know if you have any additional questions.

Yours sincerely

Hans-Joerg Fugel

Section A: Clarification on effectiveness data

Literature searching

A1. Regarding the Medline/Embase strategies reported for all sections, please clarify if this was a single search conducted simultaneously over both the Embase and Medline individual databases or was it a single search of Embase conducted on the understanding that it now contains all records from Medline.

A single search was conducted simultaneously for both Embase and Medline using the Embase.com platform. Separate searches were conducted for retrieving Medline In-Process records and this was done through the Pubmed.com platform.

A2. Please provide search dates for the conference searches for all sections and provide details of any search strategies used.

Conference searches were conducted in June 2016. Details of the search terms are listed in Table 1.

Table 1: Conference search terms

	Search terms
Disease terms	obes, adipos, overnutrition, hyperphagia, appetite, satiety, weight reduction, overweight, body mass, BMI
Intervention terms	mysimba, naltrexone, bupropion, contrave, orlistat, xenical, alli, beacita, tetrahydrolipstatin, 'mysimba'; 'naltrexone-bupropion' OR 'naltrexone/bupropion' OR 'naltrexone / bupropion' OR (naltrexone NEAR/5 bupropion) OR 'bupropion/naltrexone' OR 'bupropion / naltrexone' OR 'schembl15633271' OR 'schembl-15633271' OR 'schembl 15633271'; 'contrave'; 'orlistat' OR 'xenical' OR 'alli' OR beacita; tetrahydrolipstatin OR 'ro 18 0647' OR 'ro 18-0647' OR 'ro 180647' OR 'ro18647' OR '96829 58 2'

A3. The ERG are concerned by the use of study design filters in the Cochrane Library searches of CDSR, DARE and CENTRAL listed in Appendix 2. We feel that this is an overly restrictive approach given that these resources are already filtered by study design. We reran your strategy and noted that the combination of Lines #11 and #17 when limited to CENTRAL brought back 338 records, however the additional study design filters which appeared in your strategy reduced this to between 268-273 depending on whether you use your limit code of "and CCRCT" (line #92) or Wiley's own Trials limit. Please rescreen these results missed by your study design filters and confirm whether any records meeting your inclusion criteria were missed by this approach.

Searches were conducted again by applying the CENTRAL limit in the Cochrane Library instead of using the study design filters, as was done originally. This found only five additional unique papers from which three were deemed relevant. However, these three potentially relevant studies were published after June 2016, when the original searches were conducted. As such, no additional studies were included from this approach.

A4. Section 4.12 talks about adverse events, however no mention is made of how this information was identified and no searches were reported. Please confirm which searches were used to inform this section. If the searches reported in Appendix 2 were used please confirm if all of the results retrieved were screened for adverse events. If additional searches were used, please provide full details.

Furthermore, please confirm that only RCTs were included, even for adverse events.

No additional searches to those reported in Section 4.1 and Appendix 2 were conducted to identify adverse event (AE) data, but results retrieved were screened for AEs.

Both RCTs (randomised controlled trials) and non-RCTs were identified through SLR (systematic literature review), and screened for AEs. However, non-RCT evidence was not formally considered as part of comparative safety assessments as RCT data were available for the intervention and comparators of interest to the decision problem. This included longer-term safety data to that available from the pivotal trial programme.

A5. Please confirm that the Econlit search was carried out on the EBSCO platform as stated in Table 22, Appendix 15. It is our understanding that EBSCO host does not support the search of MH as a field in Econlit as shown in your strategy.

Econlit searches were carried out through the EBSCO platform only; however, the MH search functionality was incorrectly presented in Table 22. The corrected search strategy is presented in Table 2.

Table 2: Search strategy for cost-effectiveness studies in Econlit

S. No.	Query	Search Options	Hits
S1	SU "obesity"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	639,193
S2	SU "morbid obesity"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	3,251
S3	SU "abdominal obesity"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	1,246
S4	SU "overnutrition" OR SU "hyperphagia" OR SU "appetite" OR SU "satiety"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	45,156
S5	SU "weight reduction"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	2,965
S6	TI (adipos* OR obes*) OR AB (adipos* OR obes*)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	1,866,669
S7	TI (overweight* OR (over N3 weight*) OR "over-weight" OR overeating OR "over-eating" OR "over eating") OR AB (overweight* OR (over N3 weight*) OR "over-weight" OR overeating OR "over-eating" OR "over eating")	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	607,855
S8	TI (weight N3 (reduc* OR decreas* OR los* OR control* OR gain* OR manage* OR maint* OR watch)) OR AB (weight N3 (reduc* OR decreas* OR los* OR control* OR gain* OR manage* OR maint* OR watch))	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	1,462,983
S9	TI (("body mass" N1 ind*) OR bmi OR "waist hip ratio" OR whr OR "skinfold thickness" OR "waist circumference" OR "body fat" OR "fat mass" OR "body weight") OR AB (("body mass" N1 ind*) OR bmi OR "waist hip ratio" OR whr OR "skinfold thickness" OR "waist circumference" OR "body fat" OR "fat mass" OR "body weight")	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	1,367,353

S. No.	Query	Search Options	Hits
S10	SU "body weights and measures"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	5,489
S11	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	4,432,903
S12	TI mysimba OR AB mysimba	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	80
S13	TI ("naltrexone-bupropion" OR "naltrexone/bupropion" OR "naltrexone / bupropion" OR (naltrexone N5 bupropion) OR "bupropion/naltrexone" OR "bupropion / naltrexone" OR "schembl15633271" OR "schembl-15633271" OR "schembl 15633271") OR AB ("naltrexone-bupropion" OR "naltrexone/bupropion" OR "naltrexone / bupropion" OR (naltrexone N5 bupropion) OR "bupropion/naltrexone" OR "bupropion / naltrexone" OR "schembl15633271" OR "schembl-15633271" OR "schembl 15633271")	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	931
S14	TI contrave AND AB contrave	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	662
S15	TI (orlistat OR xenical OR alli OR beacita) OR AB (orlistat OR xenical OR alli OR beacita)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	140,031
S16	TI (tetrahydrolipstatin OR "ro 18 0647" OR "ro 18-0647" OR "ro 180647" OR "ro18647" OR "96829 58 2") OR AB (tetrahydrolipstatin OR "ro 18 0647" OR "ro 18-0647" OR "ro 180647" OR "ro18647" OR "96829 58 2")	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	716
S17	TI (sibutramine OR sibutramin* OR arcalion) OR AB (sibutramine OR sibutramin* OR arcalion)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	8,249
S18	TI ("bts 54 524" OR "bts 54524" OR "bts54524") OR AB ("bts 54	Expanders - Also search within the full text of the articles	11

S. No.	Query	Search Options	Hits
	524" OR "bts 54524" OR "bts54524")	Search modes - Find all my search terms	
S19	TI (reductil OR medaria OR meridia OR "106650 56 0") OR AB (reductil OR medaria OR meridia OR "106650 56 0")	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	4,560
S20	TI (rimonabant OR acomplia OR zimulti) OR AB (rimonabant OR acomplia OR zimulti)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	8,753
S21	TI ("sr 141716" OR "sr141716" OR "sr 141716a" OR "sr141716a" OR "158681 13 1") OR AB ("sr 141716" OR "sr141716" OR "sr 141716a" OR "sr141716a" OR "158681 13 1")	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	5,304
S22	TI (lorcaserin OR belviq) OR AB (lorcaserin OR belviq)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	3,693
S23	TI ((phentermine AND topiramate) OR qsymia) OR AB ((phentermine AND topiramate) OR qsymia)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	2,023
S24	TI (liraglutide OR saxenda OR victoza OR nn2211) OR AB (liraglutide OR saxenda OR victoza OR nn2211)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	9,437
S25	S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	178,999
S26	SU "Economics"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	2,668,302
S27	SU "Costs and Cost Analysis"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	197,274
S28	SU "Cost Allocation"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	5,472
S29	SU "Cost-Benefit Analysis"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	106,275

S. No.	Query	Search Options	Hits
S30	SU "Cost Control"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	151,222
S31	SU "Cost Savings"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	12,829
S32	SU "Cost of Illness"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	21,658
S33	SU "Cost Sharing"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	4,817
S34	SU "Deductibles and Coinsurance"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	1,547
S35	SU "Medical Savings Accounts"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	4,151
S36	SU "Health Care Costs"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	56,360
S37	SU "Direct Service Costs"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	1,119
S38	SU "Drug Costs"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	14,084
S39	SU "Employer Health Costs"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	1,093
S40	SU "Hospital Costs"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	12,523
S41	SU "Health Expenditures"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	17,724

S. No.	Query	Search Options	Hits
S42	SU "Capital Expenditures"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	5,375
S43	SU "Value of Life"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	6,200
S44	SU "Economics, Medical"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	14,505
S45	SU "Economics, Hospital"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	21,491
S46	SU "Economics, Nursing"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	3,964
S47	SU "Economics, Pharmaceutical"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	2,718
S48	SU "Budgets"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	142,096
S49	SU "Fees and Charges"	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	29,166
S50	TI (low N1 costs) OR AB (low N1 costs)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	734,885
S51	TI (high N1 costs) OR AB (high N1 costs)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	408,389
S52	TI (healthcare N1 cost*) OR AB (healthcare N1 cost*)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	65,619
S53	TI ((fiscal OR funding OR financial OR finance)) OR AB ((fiscal OR funding OR financial OR finance))	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	45,553,073

S. No.	Query	Search Options	Hits
S54	TI (cost N1 estimate*) OR AB (cost N1 estimate*)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	816,085
S55	TI (cost N1 variable*) OR AB (cost N1 variable*)	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	14,278
S56	TX unit N1 cost*	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	891,210
S57	TI (economic* OR pharmacoeconomic* OR price* OR pricing OR cea OR cua OR markov OR (decision N2 tree*) OR (decision N2 analysis*) OR (monte N1 carlo)) OR AB (economic* OR pharmacoeconomic* OR price* OR pricing OR cea OR cua OR markov OR (decision N2 tree*) OR (decision N2 analysis*) OR (monte N1 carlo))	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	46,837,666
S58	S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	93,291,093
S59	S11 AND S25 AND S58	Expanders - Also search within the full text of the articles Search modes - Find all my search terms	4,773
S60	S11 AND S25 AND S58	Expanders - Also search within the full text of the articles Search modes - Find all my search terms Limiters - Date Published: 20060101-20160531	649
S61	S11 AND S25 AND S58 Source: Econlit	Expanders - Also search within the full text of the articles Search modes - Find all my search terms Limiters - Date Published: 20060101-20160531	2

A6. The Econlit search in Appendix 15 appears to contain an error in the line numbers being combined in lines S60 and S61. The line above (S59) has the combination

"S11 AND S25 AND S58" which appears to be correct; however the following two lines have the combination "S11 AND S22 AND S58. Line S25 is a combination of all listed interventions where line S22 is for "TI (lorcaserin OR belviq) OR AB (lorcaserin OR belviq)". Please confirm whether this was an error in reporting or one which occurred during the search.

An error was made when reporting the search. The line S60 AND S61 have been updated to S11 AND S25 AND S58, as presented above in Table 2.

Outcomes

A7. Body Mass Index (BMI)

a. Please justify why BMI was not evaluated as an outcome in the submission according to the NICE scope?

BMI was considered within the economic modelling, but was not explicitly provided as a clinical outcome of the four COR trials as this was not a pre-defined endpoint.

b. **Priority**: Mean BMI at baseline is provided for each of the four main naltrexone-bupropion trials. Please calculate mean BMI at week 56 (or end-of-study) using the height at baseline for all four trials (COR-I, COR-II, COR-BMOD and COR-DM).

Baseline patient BMI was calculated for all four trials (COR-I, COR-II, COR-BMOD and COR-DM) as part of inclusion criteria, however, BMI was not a primary or secondary endpoint for these clinical trials and has not previously been calculated for week 56. Orexigen is currently analysing the patient level data in order to provide the mean BMI at week 56 (or end-of-study) but this analysis will not be available until 20th February 2017.

Clinical trial data and results

A8. In section 2.2 on page 26 of the company submission an ongoing phase IV study is mentioned on the occurrence of MACE in overweight and obese patients taking NB32. Please provide bibliographical details of the trial? Please provide the protocol. Are any interim data available? If so then please provide these?

Study synopsis is provided as an attachment. No information related to the new MACE study has been published or is available on any bibliographic database as it is currently still in the planning stage.

A9. Please provide bibliographic details and the protocol of the ongoing trial evaluating safety of NB32 in patients with renal or hepatic impairment mentioned in section 2.2 on page 26?

Study synopsis are provided as an attachment. As both the renal and hepatic impairment studies are small phase I studies requested by regulatory agencies, no information related to

these studies have been published or made available on clinical study databases, such as clinicaltrials.gov.

A10. **Priority**: In section 2.3 table 6 there is a statement "Retreatment with NB32 is not routinely anticipated and thus not modelled." Please justify why patients would not be retreated with naltrexone-bupropion for any subsequent weight gain after a successful treatment with the drug?

As per the SmPC, patients who respond to treatment should stay on NB32 to continue to benefit from the medication, including improvements in weight-related comorbidities, such as hypertension, prediabetes, and diabetes. There are no data to indicate the effectiveness of retreatment with NB32 following successful treatment with NB32 and subsequent discontinuation and weight regain. If NICE thinks this is likely to happen in practice, an option for NICE is to consider that the current cost-effectiveness model assumes the same analysis for patients independent of whether they have received previous NB32 or not. Clinical rationale can inform the likelihood of retreatment success until evidence merges. If treatment effect is unlikely to diminish, given the economic analysis is extremely conservative in assuming minimal impact on downstream comorbidities, we believe subsequent treatment with NB32 following successful initial treatment can be conservatively derived from the existing results.

A11. **Priority**: In section 2.3 table 6 there is a statement 'For patients continuing treatment post 16 weeks, treatment should be continued as long as clinical benefit is observed.' The main trials are just over a year's duration.

a. Please provide the precise criteria by which treatment discontinuation was determined in the trials?

In the four COR trials, patients were free to discontinue (i.e. withdrew consent or no longer willing to participate) their participation in the study at any time and without any prejudice to further treatment. The investigator could withdraw a patient at any time because of a safety risk of AE.

The study drug may have been discontinued for any of the following reasons:

- Intercurrent illness or condition that would, in the judgement of the investigator, affect assessments of clinical status to a significant degree or put the patient at increased risk
- Unacceptable toxicity, which the investigator judged to compromise subject safety or the ability to perform study-specific procedures, or not to be in the subject's best interest
- Suicide attempt
- Seizure
- Patient requested to discontinue treatment for any reason

- Patient in the titration phase who was unable to take the study drug at the prescribed dose for more than 72 consecutive hours due to intolerable AEs
- Patients who stop the study drug for any reason for a period of 15 consecutive days or longer
- Non-compliance, as defined by failure of the subject to return for two or more consecutive study visits, or failure to adhere to 70% compliance for 2 consecutive months
- An increase in the patient's alanine transaminase (ALT) and/or aspartate transaminase (AST) of five times the upper limit of normal
- Pregnancy
- Discontinuation of the study at the request of Orexigen Therapeutics

In addition, in the COR-DM study, patients may have discontinued treatment if they required insulin therapy for >14 consecutive days.

b. Please clarify what was the percentage of participants in each of the trials that discontinued due to cessation of clinical benefit?

In the four COR trials, insufficient weight loss was defined as a lack of efficacy. The proportion of patients who discontinued due to insufficient weight loss is presented in Table 3.

Table 3: Patient discontinuations due to insufficient weight loss in the COR trials

Trial	Patients who discontinued due to insufficient weight loss, n (%)		
	Total	Treatment arm	
COR-I[1]	64 (3.7)	NB32	12 (2.1)
		NB16	12 (2.1)
		Placebo	40 (6.9)
COR-II[2]	52 (3.5)	NB32	19 (1.9)
		Placebo	33 (6.7)
COR-BMOD[3]	9 (1.1)	NB32 + BMOD	3 (0.5)
		Placebo + BMOD	6 (2.9)
COR-DM[4]	11 (2.2)	NB32	5 (1.5)
		Placebo	6 (3.5)

Key: BMOD, behavioural modification; NB16, naltrexone 16mg plus bupropion; NB32, naltrexone 32mg plus bupropion.

c. Please provide the precise criteria by which treatment discontinuation would be determined in clinical practice?

As a result of pooled, *post-hoc* analyses of the COR trials that showed a strong relationship between early weight loss and clinically meaningful longer term weight loss, the license terms for NB32 include more prescriptive discontinuation rules. As stated in the summary of product characteristics, the need for continued treatment should be evaluated after 16 weeks and treatment should be discontinued if patients have not lost at least 5% of their initial body weight. Professor Wilding supported the ongoing criterion for treatment continuation being maintenance of a loss of at 5% body weight from baseline. He also stated that current guidelines (for orlistat) recommend a review of the need for ongoing treatment at 1 year.[5]

A12. **Priority**: Please justify not including standard management as an intervention in the review eligibility criteria and the searches? Currently standard care is considered only as a comparator to orlistat or naltrexone-bupropion. Therefore, any studies comparing behavioural interventions with no treatment (or other behavioural interventions) are excluded. However, these are relevant according to the scope.

The anticipated positioning of NB32 in the treatment pathway is for patients eligible for pharmacological treatment (alongside standard management), therefore if we had specifically searched for standard management publications we would likely have introduced a large amount of heterogeneity and do not think that the evidence is for this patient population is directly relevant to the decision problem. In addition to this, standard management, and varying types of standard management reflective of what is seen in clinical practice, is available directly through head-to-head trial data for both orlistat and NB32. Defining and consolidating the RCT data alone had its own challenges in terms of the variability of standard management, and the head-to-head RCT evidence is directly relevant to the decision problem. Introducing further heterogeneity when extensive existing high level evidence is already available for standard management should not be considered in this case.

Further, from a practical perspective, current standard of care was not in the pre-referral draft scope and was included as “standard management without naltrexone-bupropion” at the post-referral scope stage (September 2016) which impacted the search criteria for the clinical systematic literature reviews (conducted in June 2016). Given the scope is standard management without NB32, the most relevant evidence for the decision problem has been presented.

A13. Please clarify why non-RCTs were eligible for the review but not considered further, not even for adverse events?

Non-RCT evidence was not formally considered as part of comparative efficacy, comparative safety or cost-effectiveness assessments as RCT data were available for the intervention

and comparators of interest to the decision problem. As noted in response to A4, RCT data included longer-term efficacy and safety data to that available from the pivotal trial programme. In light of its completeness, non-RCT data was not deemed pertinent to the decision problem.

A14. Table 10, Eligibility criteria for trials to be included in the systematic review, (CS, page 44) mentions that 'Studies published in non-English languages were flagged'. Please explain what is meant by that and please explain what was done with these studies.

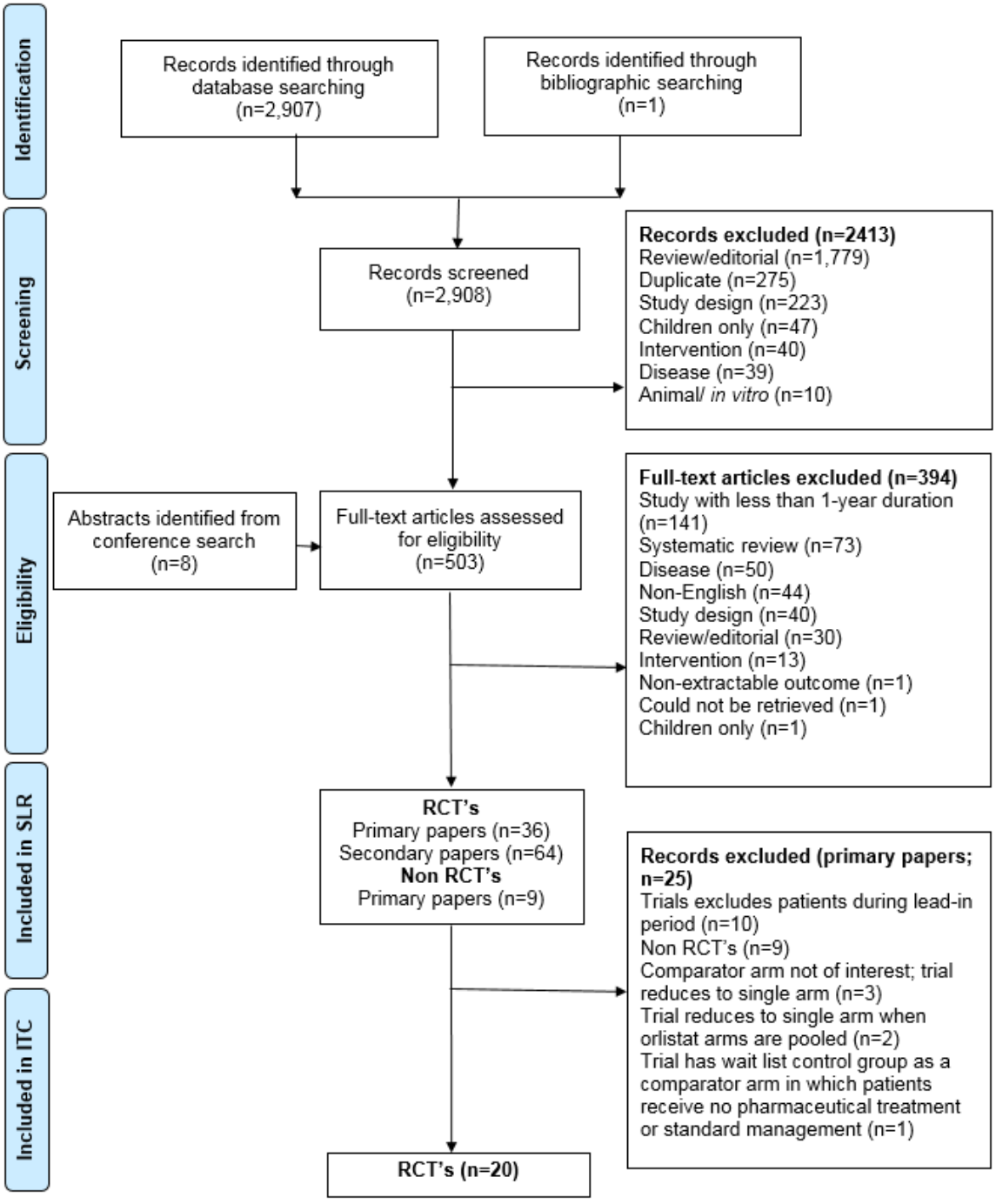
Non-English language studies were to be included if sufficient evidence from English language articles was not available. In light of the completeness of English language RCTs, all non-English language studies were excluded.

A15. Please extend the flow chart in Figure 2 of the submission to illustrate the studies considered in the direct meta-analysis and those forming part of the indirect comparison with reasons for exclusion?

The extended flow chart is presented in Figure 1.

.

Figure 1: PRISMA flow diagram of the clinical effectiveness literature search process (May 2016)



Key: ITC, indirect treatment comparison; n, number of studies; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-analyses; RCT, randomised controlled trial. SLR, systematic literature review.

A16. Please clarify how many, if any, of the patients in the COR trials and NB-CVOT had previously received treatment with orlistat?

In all four COR trials, the exclusion criteria includes “treatment with any anorectic or weight loss agent”, although in the COR-I study, one patient received prior treatment with orlistat. The one patient who received orlistat was a protocol violation

In the NB-CVOT study, patients were prohibited from taking additional weight loss medication, although one patient was recorded as having received orlistat at screening, Year 1 and Year 2.

A17. Please provide a summary table of percentages of patients in the COR trials and NB-CVOT who are both overweight according to the NICE scope (≥ 27 kg/m^2 to < 30 kg/m^2) and with one or more weight-related co-morbidities?

The % breakdown of patients in each weight class for the four studies are outlined below:

Study 301 (COR-1)

Obesity class	Placebo	NB16	NB32
BMI<30kg/m^2	0.9%	2.8%	31.1%
BMI ≥30 and ≤35 kg/m^2	37.3%	37.7%	38.4%
BMI ≥35 and <40 kg/m^2	39.4%	34.6%	35.0%
BMI ≥ 40 kg/m^2	22.4%	24.9%	23.5%

Study 302 (COR-BMOD)

Obesity class	Placebo	NB32
BMI<30kg/m^2	0.5%	1.4%
BMI ≥30 and ≤35 kg/m^2	31.7%	35.0%
BMI ≥35 and <40 kg/m^2	39.1%	38.9%
BMI ≥ 40 kg/m^2	28.7%	24.7%

Study 303 (COR-2)

Obesity class	Placebo	NB32
BMI<30kg/m^2	2.8%	2.5%
BMI ≥30 and ≤35 kg/m^2	37.6%	39.8%
BMI ≥35 and <40 kg/m^2	38.6%	31.6%
BMI ≥ 40 kg/m^2	21.0%	26.2%

Study 304 (COR-DM)

Obesity class	Placebo	NB32
BMI<30kg/m^2	6.5%	5.4%
BMI ≥30 and ≤35 kg/m^2	28.8%	33.1%
BMI ≥35 and <40 kg/m^2	37.6%	32.8%
BMI ≥ 40 kg/m^2	27.1%	28.7%

CVOT study

Obesity class	Placebo	NB32
BMI<30kg/m^2	7%	6.7%
BMI ≥30 and ≤35 kg/m^2	31.6%	31.3%
BMI ≥35 and <40 kg/m^2	38.6%	38.0%
BMI ≥ 40 kg/m^2	30.3%	28.8%

A18. **Priority**: Please describe in more detail the components of standard care in the four COR trials? Please include summary statistics of number of contacts with each type of health care professional as well as any specific instructions to exercise or to attend a weight loss club. How was consistency of standard care between centres within a trial assured?

In the COR-I and COR-II studies, all patients received ancillary therapy at baseline and Weeks 12, 24, 26 and 48. Patients were instructed to follow a hypocaloric diet representing a deficit of 500 kcal per day based on the World Health Organization algorithm for calculating resting metabolic rate. Adjusted body weight was used to calculate energy needs because subjects were 120% greater than ideal body weight. Subjects received written instructions on behavioural modification techniques. Patients were encouraged to increase physical activity, with a prescription for walking starting with at least 10 minutes on most days of the week, and increasing this gradually to 30 minutes on most days of the week throughout the study. They were encouraged to lose weight and maintain weight loss, and were encouraged to follow the prescribed programme (as described). Participation in any other weight loss programme was not permitted. The use of meal replacements (such as Slim Fast® or Weight Watchers®) was discouraged, but occasional use did not necessitate withdrawal from the study. The prescribed exercise could be performed in a gymnasium or health club.

In the COR-BMOD study, all patients were to participate in an intensive behaviour modification program that included three components: dietary instruction, closed group sessions, and prescribed exercise. Behaviour modification consisted of group meetings (10 to 20 patients per session) lasting 90 minutes (including weigh-in) weekly for the first 16 weeks, every other week for the next 12 weeks and monthly thereafter for up to 28 sessions. They included instructions to consume a balanced deficit diet and to increase to 180 min/week of planned, moderately vigorous, physical activity. Dietary instructions were provided at baseline (Day 1). Patients began closed group sessions no later than 4 weeks after randomisation.

In the COR-DM study, all patients received ancillary weight loss therapy at baseline and Weeks 4, 16, 28, and 40. Ancillary therapy consisted of diet instruction, behaviour

modification advice and physical activity suggestions. Patients were instructed to follow a hypocaloric diet representing a deficit of 500 kcal/day based on the World Health Organization's algorithm for calculating resting metabolic rate. Adjusted body weight was used to calculate energy needs because subjects were 120% greater than ideal body weight. Patients received behavioural modification advice, including written instructions. Dietary counselling was conducted in accordance with the American Diabetes Association and American Dietetic Association guidelines for counselling diabetics. "Exchange Lists for Weight Management, 2nd edition" booklets were provided to trial participants to facilitate adherence to prescribed dietary regime. Patients were encouraged to increase physical activity, with a prescription for walking at least 30 minutes three times per week. Patients were encouraged to follow the prescribed programme. Participation in any other organised weight loss programme was not permitted. The use of meal replacements (such as Slim Fast® or Weight Watchers®) was discouraged, but occasional use despite contrary instructions did not necessitate withdrawal from the study. The prescribed exercise could be performed in a gymnasium.

Compliance was only measured for study medication with no check of compliance for the diet and exercise regimens. No summary statistics for the placebo group were captured in any of the clinical trials.

A19. ITT analysis

a. Please justify the use of a modified ITT analysis in the COR trials?

The purpose of using the modified ITT was to be able to compare patients who have received at least one dose of NB with patients treated with placebo, requiringat least one on-treatment post-baseline weight recorded. The patients who were included in the modified ITT analysis had to meet the following 3 criteria:

- A baseline body weight was recorded.
- Patient was randomized
- A post-baseline body weight was recorded while patient was on treatment.

b. **Priority**: Please provide all clinical effectiveness outcomes from the NB trials used in the economic model based on two ITT populations: "ITT with the Weight regain imputation method" and "ITT with Baseline-carried forward analysis". In other words, please provide all data for NB32 and SM as reported in CS Tables 56-58 (proportion of responders, average weight loss at the assessment moments) for these two ITT populations. Please also provide data on treatment discontinuation (before, between and after the two assessments moments) based on these two populations.

The complete set of clinical effectiveness outcomes derived from the two requested populations (that is, "intention-to-treat (ITT) with the weight regain imputation method" and

"ITT with baseline-carried forward analysis" [BOCF]) are not immediately available for all four COR-trials at the present time. Table 4 presents the data available for the clinical effectiveness outcomes for the requested populations.

Table 4: Data availability for the clinical effectiveness outcomes for the requested populations

Trial	Placebo corrected LS mean % of patients achieving at least 5% weight loss (95% CI)		Placebo corrected LS mean % change from baseline at 56 weeks (SE)	
	ITT-BOCF	ITT-WRIM*	ITT-BOCF	ITT-WRIM*
COR-I	30.9 (27.1, 34.6)	34.8 (31.0, 38.7)	-4.0(0.3)	-4.6(0.3)
COR-II	45.5(41.5, 9.5)	51.4 (47.4, 55.5)	-6.4(0.4)	-7.3(0.4)
COR-DM	35.1 (32.0, 38.2)	38.4 (35.2, 41.5)	-4.4(0.2)	-4.9(0.2)
COR-BMOD	28.1 (23.3, 32.9)	31.0(26.1, 36.0)	-3.1(0.3)	-3.5(0.3)

Key: BMOD, behaviour modification; BOCF, baseline observation carried forward; CI, confidence interval; COR, Contrave® Obesity Research; DM, diabetes mellitus; ITT, intention-to-treat; LS, least squares; SE, standard error; WRIM, weight regain imputation method.
Notes: WRIM assumes patients regain 0.3kg per month following study withdrawal.

Whilst additional analyses of BOCF and WRIM populations may become available at a later date, it is important to highlight the differences between the two requested populations and the modified ITT (mITT) population used to inform both the clinical evidence in the manufacturer's submission and the *de novo* economic model. A total of 2,393 patients formed the ITT population for NB32 across the COR trial programme. Of these, 2,043 patients formed the mITT population (i.e. approximately 85% of the total randomised population).

The mITT population was defined in the manufacturer's submission as patients who had at least one post-baseline weight measurement obtained while the patient was still taking study medication, with missing data imputed using the last observation carried forward (LOCF) method. To derive weight loss outcomes for patients beyond 16 weeks, it was required to establish the cohort of patients that responded at 16 weeks. Regardless of population utilised, the subset of patients that responded to treatment at Week 16 is the same. As such, the only weight loss outcomes required for the model that could utilise the ITT populations are those at the Week 16 assessment.

However, within the economic analysis, weight loss outcomes were separated by those who respond to treatment and those who do not, with a randomly sampled number utilised to determine whether the patient is a responder or a non-responder. By utilising weight loss outcomes for patients with no further observations from baseline (as is implied by considering the ITT populations over the mITT population), the proportion of primary assessment non-responders will be over-estimated as the analysis will automatically assign all patients with no further measurements as non-responsive.

The use of BOCF to impute missing data may result in further overestimation of the number of non-responders, as a patient that discontinues from the study post baseline is also assumed to have had no change in weight; a patient that discontinues towards the end of the study would therefore be assumed to have received no treatment effect, which is clearly unlikely.

Although data imputation using LOCF avoids this issue, it is acknowledged that patients are likely to regain weight post discontinuation of treatment (that is, standard management and adjunctive therapy). This has been considered in the economic model, which applies a linear regain period of 3 years that commences upon a patient discontinuing treatment. Implementation of weight-regain using this method allows appropriate assumptions to be applied across all patients (i.e. regardless of whether a patient has missing data or not).

The application of weight regain within the model is likely to provide a more accurate estimate of a patient's regain in weight than the "ITT with the weight regain imputation method", which assumes that, regardless of a patient's baseline and current weight, their weight would increase at a rate of 0.3kg per month until they return to their baseline weight. Use of a regain rate per month can result in very large/very small regain periods (e.g. if a patient loses 10kg and begins to regain weight, the patient would never regain their weight fully as this would take 30 years).

The use of either of the requested ITT populations is therefore likely to result in bias against NB32. Furthermore, this bias is likely to be extended if either population are considered within the indirect treatment comparison as data for the requested populations are not available within the orlistat trials. In order to avoid implementing systematic errors of this nature within the model the mITT population was preferred.

Within the time to treatment discontinuation (TTD) analysis, the safety population (i.e. not the mITT population) was utilised for patients between t=0 and t=16. As such, untreated patients (who would feature in ITT analyses) would be censored automatically at t=0. The resultant Kaplan–Meier (KM) function for TTD would be expected to be the same, other than having a slightly larger number of risk at t=0. The difference between the number at risk at the beginning of the KM for the ITT and safety populations would be the number of patients immediately censored at t=0, hence the resultant functions would be equivalent. Beyond t=16 weeks, similar logic to the weight loss outcomes applies (i.e. the population required must be a subset of the mITT population in order to distinguish between responders and non-responders).

In summary, the ITT populations are generally not applicable to the *de novo* model, and inclusion of these patients would lead to an overestimation of the proportion of patients who fail to respond at 16 weeks, and no difference to the estimation of TTD.

c. **Priority**: Please provide all clinical effectiveness outcomes from the COR-BMOD trial for the control arm (intensive behaviour modification only) if used in the economic model based on two ITT populations: "ITT with the Weight regain imputation method" and "ITT with Baseline-carried forward analysis". In other words, please provide all data from the control arm of COR-BMOD as reported in CS Tables 56-58 (proportion of responders, average weight loss at the assessment moments) for these two ITT populations. Please also provide data on treatment discontinuation (before, between and after the two assessments moments) based on these two populations.

See answer to Question 19b.

A20. Please justify why only 5% reduction in weight and mean % weight change from baseline at 1 year were chosen as outcomes for the meta-analysis?

Regarding the ultimate application of results from the meta-analysis to the *de novo* economic model, outcomes of 5% reduction in weight and mean % weight change from baseline were the only outcomes required from the meta-analysis. The outcome of 5% reduction in weight from baseline was incorporated as per the European Medicines Agency (EMA) licence and associated treatment stopping rules; whereas the mean % weight change from baseline was incorporated to account for the overarching treatment effect of each regimen. Meta-analysed results for alternate outcomes were not required for the *de novo* model, and were therefore not produced.

A21. Please provide four clinical effectiveness outcomes (Mean % weight change from baseline at 1 year; 5% reduction in weight at 1 year from baseline; Change in waist circumference (cm) at 1 year; and Proportion of patients with ≥10% decrease in body weight at 1 year (the 1-year time point ranging from 52 to 57 weeks)) from four NB32 trials (COR-I, COR-II, COR-BMOD and COR-DM) based on two ITT populations: "ITT with the Weight regain imputation method" and "ITT with Baseline-carried forward analysis". And please provide the same meta-analyses results based on these two ITT populations as reported in chapter 4.9 in the CS.

In addition to the available data for the current clinical effectiveness outcomes presented in Table 4, Table 5 presents the data availability for change in waist circumference (cm) at 1 year and proportion of patients with ≥10% decrease in body weight at 1 year (the 1-year time point ranging from 52 to 57 weeks).

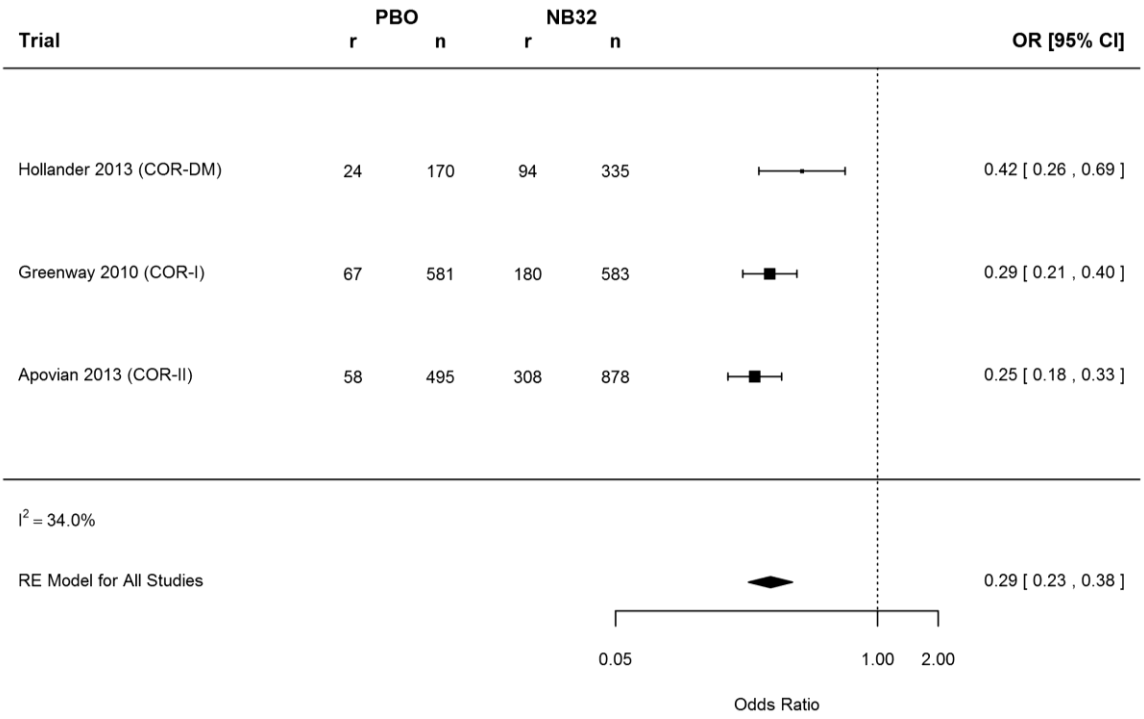
Table 5: Data availability for requested additional clinical effectiveness outcomes

Trial	At least 10% reduction in weight at 1 year			Waist circumference (cm) CFB at 1 year		
	mITT-LOCF	ITT-BOCF	ITT-WRIM	mITT-LOCF	ITT-BOCF	ITT-WRIM
COR-I	Y	Y	N	Y	N	N
COR-II	Y	Y	N	Y	N	N
COR-DM	Y	N	N	Y	N	N
COR-BMOD	Y	N	N	Y	N	N

Key: BMOD, behaviour modification; BOCF, baseline observation carried forward; CFB, change from baseline; COR, Contrave® Obesity Research; DM, diabetes mellitus; ITT, intention-to-treat; LOCF, last observation carried forward; mITT, modified intention to treat; WRIM, weight regain imputation method.
Notes: WRIM assumes patients regain 0.3 kg per month following study withdrawal.

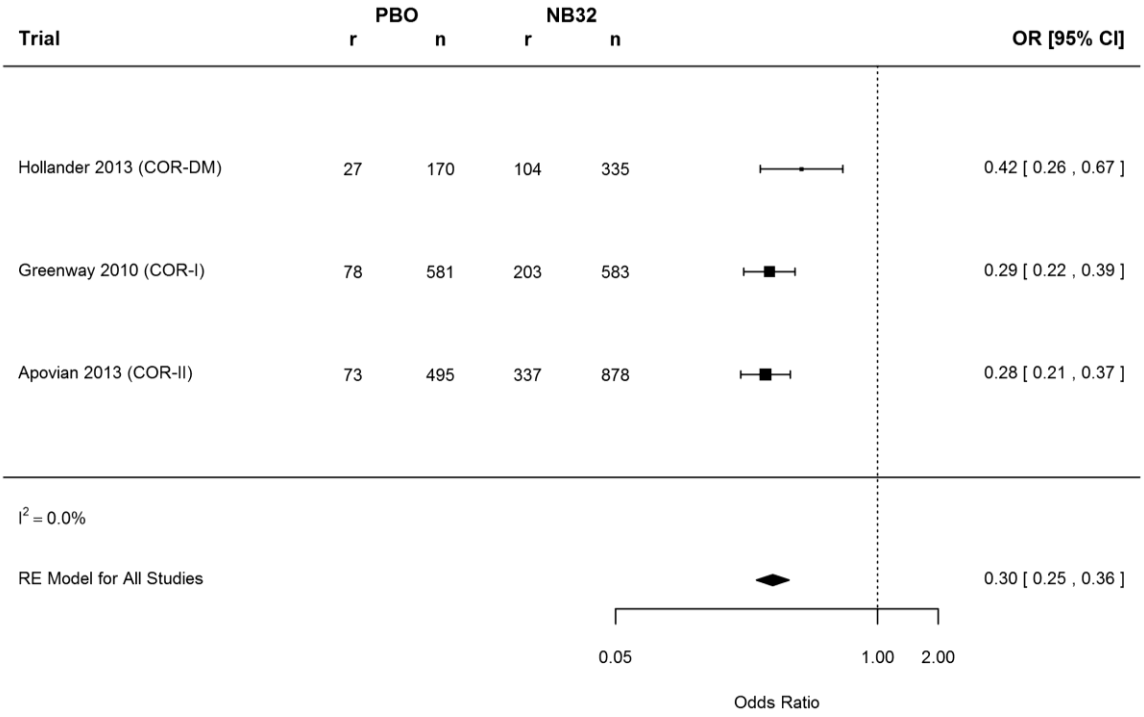
Figure 2 to Figure 8 presents the results of the direct meta-analyses for the available data.

Figure 2: Forest plot of ≥5% reduction in weight for placebo versus NB32 (ITT with BOCF)



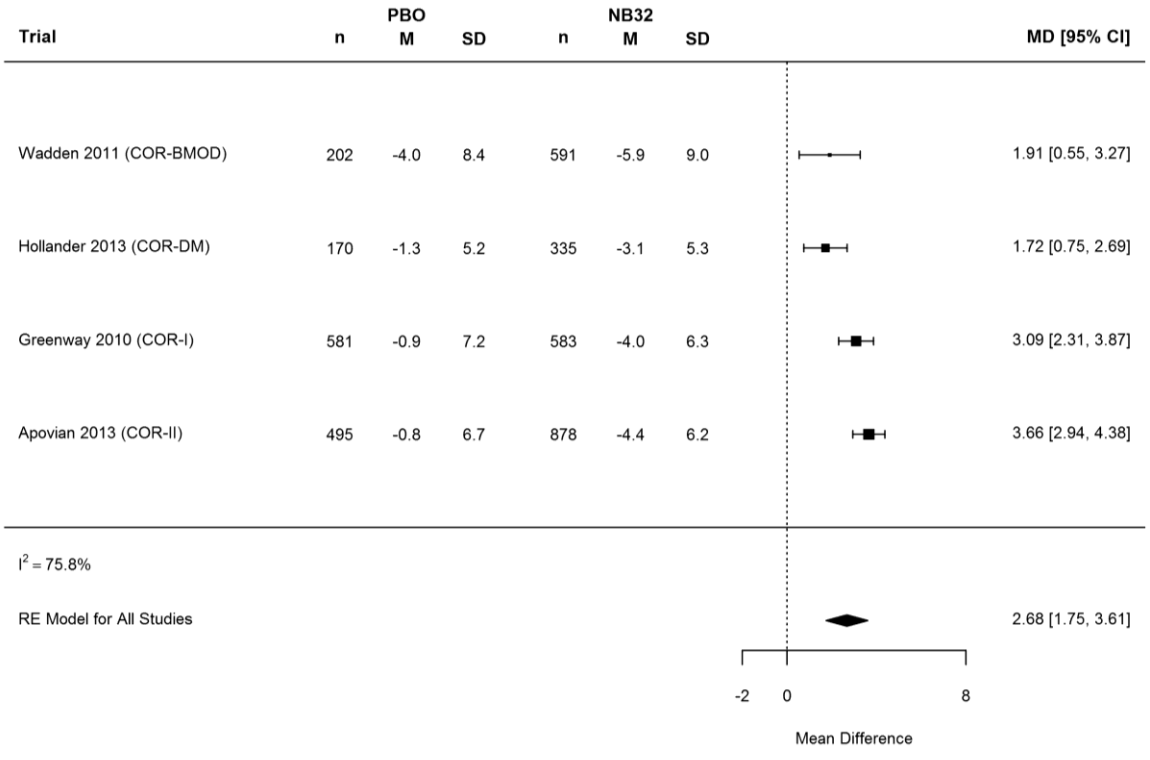
Key: BMOD, intensive behaviour modification; BOCF, baseline observation carried forward; CI, confidence interval; COR, Contrave® obesity research; DM, diabetes mellitus; ITT, intention-to-treat; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; r, number of patients achieving ≥5% reduction in weight; RE, random effects.
Notes: an odds ratio < 1 favours NB32; data were not available for the COR-BMOD study; COR-DM data from CSR.

Figure 3: Forest plot of ≥5% reduction in weight for placebo versus NB32 (ITT with WRIM*)



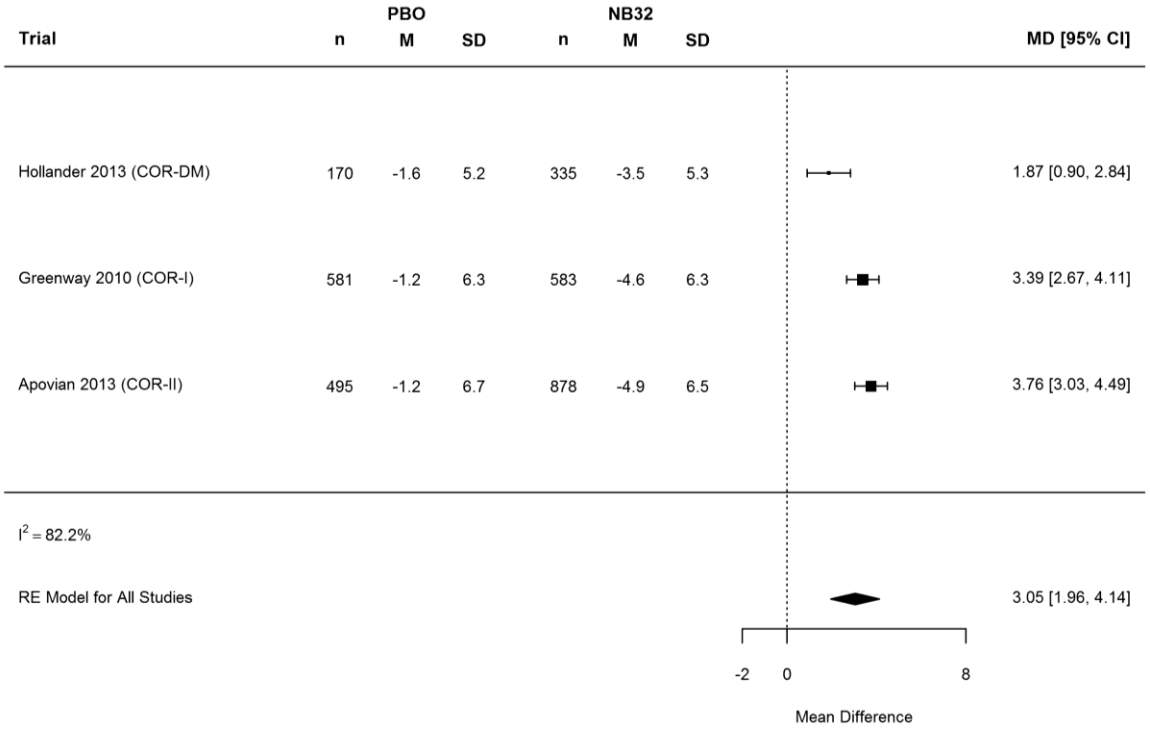
Key: BMOD, intensive behaviour modification; BOCF, baseline observation carried forward; CI, confidence interval; COR, Contrave® obesity research; DM, diabetes mellitus; ITT, intention-to-treat; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; r, number of patients achieving ≥5% reduction in weight; RE, random effects, WRIM, weight regain imputation method.
Notes: *, WRIM assumes patients regain 0.3 kg per month following study withdrawal; an odds ratio < 1 favours NB32; data were not available for the COR-BMOD study; COR-I, COR-II and COR-DM data from CSR.

Figure 4: Forest plot for % weight CFB for placebo versus NB32 (ITT with BOCF)



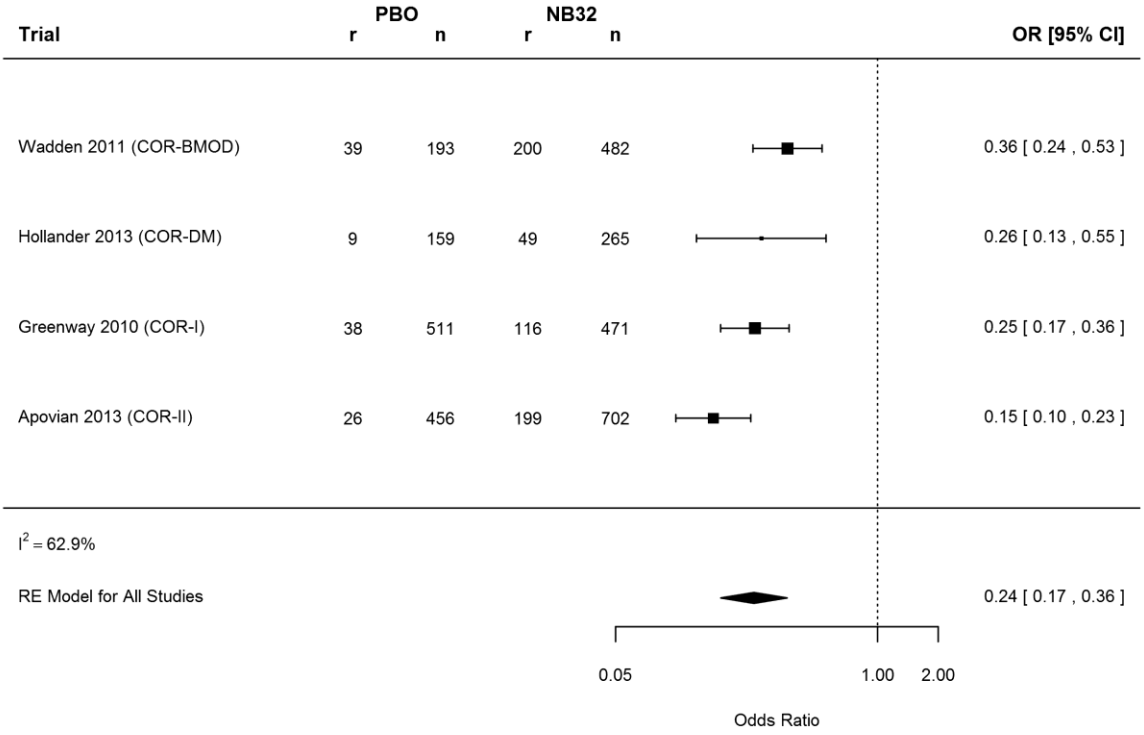
Key: BMOD, behaviour modification; BOCF, baseline observation carried forward; CFB; change from baseline; COR, Contrave® obesity research; DM, diabetes mellitus; ITT, intention-to-treat; M, mean; MD, mean difference; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; SD, standard deviation; RE, random effects.
Notes: A MD > 0 favours NB32; all data from CSR.

Figure 5: Forest plot for % weight CFB for placebo versus NB32 (ITT with WRIM*)



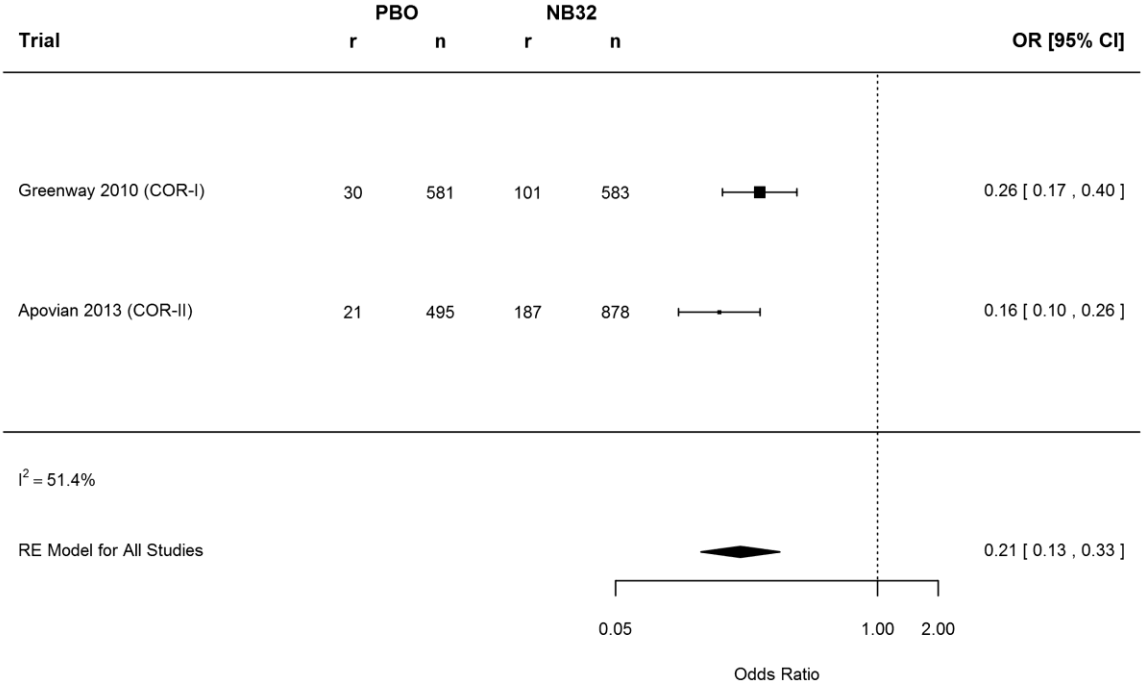
Key: BMOD, behaviour modification; BOCF, baseline observation carried forward; CFB; change from baseline; COR, Contrave® obesity research; DM, diabetes mellitus; ITT, intention-to-treat; M, mean; MD, mean difference; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; SD, standard deviation; RE, random effects.
Notes: *, WRIM assumes patients regain 0.3 kg per month following study withdrawal; a MD > 0 favours NB32; data were not available for the COR-BMOD study; all data from CSR.

Figure 6: Forest plot of ≥10% reduction in weight for placebo versus NB32 (mITT with LOCF)



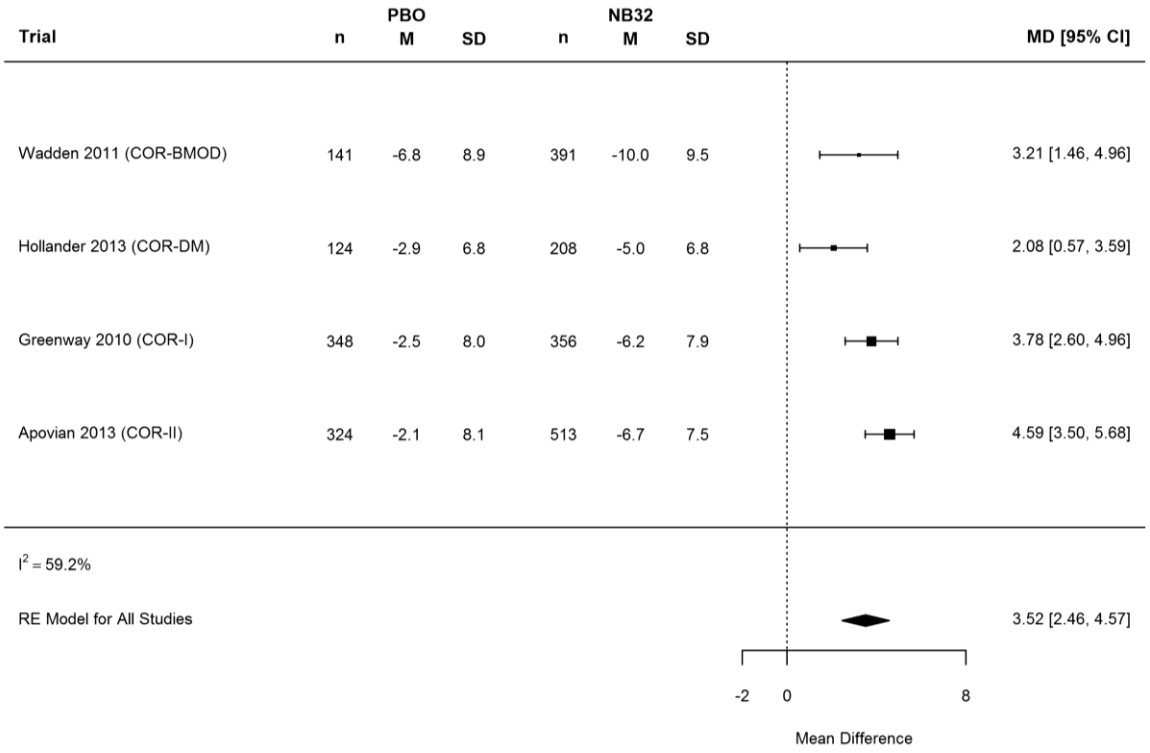
Key: BMOD, intensive behaviour modification; CI, confidence interval; COR, Contrave® obesity research; DM, diabetes mellitus; LOCF, last observation carried forward mITT, modified intention-to-treat; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; r, number of patients achieving ≥5% reduction in weight; RE, random effects.
Notes: An OR < 1 favours NB32.

Figure 7: Forest plot of ≥10% reduction in weight for placebo versus NB32 (ITT with BOCF)



Key: BMOD, intensive behaviour modification; BOCF, baseline observation carried forward; CI, confidence interval; COR, Contrave® obesity research; DM, diabetes mellitus; ITT, intention-to-treat; LOCF, last observation carried forward; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; r, number of patients achieving ≥5% reduction in weight; RE, random effects.
Notes: An OR < 1 favours NB32; Data were not available for COR-DM or COR-BMOD.

Figure 8: Forest plot for waist circumference CFB for placebo versus NB32 (mITT with LOCF)



Key: BMOD, behaviour modification; BOCF, baseline observation carried forward; CFB; change from baseline; COR, Contrave® obesity research; DM, diabetes mellitus; ITT, intention-to-treat; M, mean; MD, mean difference; n, number of patients; NB32; naltrexone 32mg plus bupropion; OR, odds ratio; PBO, placebo; SD, standard deviation; RE, random effects.
Notes: A MD > 0 favours NB32; Data from CSR.

Section B: Clarification on cost-effectiveness data

Treatment effectiveness

B1. **Priority**: Ara et al.[6] state that 'although there was a wide variation in the modelling approaches and evidence used in the studies, the variable reported to have the largest effect on the results in the majority of the models was the period of weight regain modelled.'[6] Several assumptions for weight regain in the CS base-case were discrepant with assumptions from the base-case analysis by Ara et al.[6] (see CS Table 53)

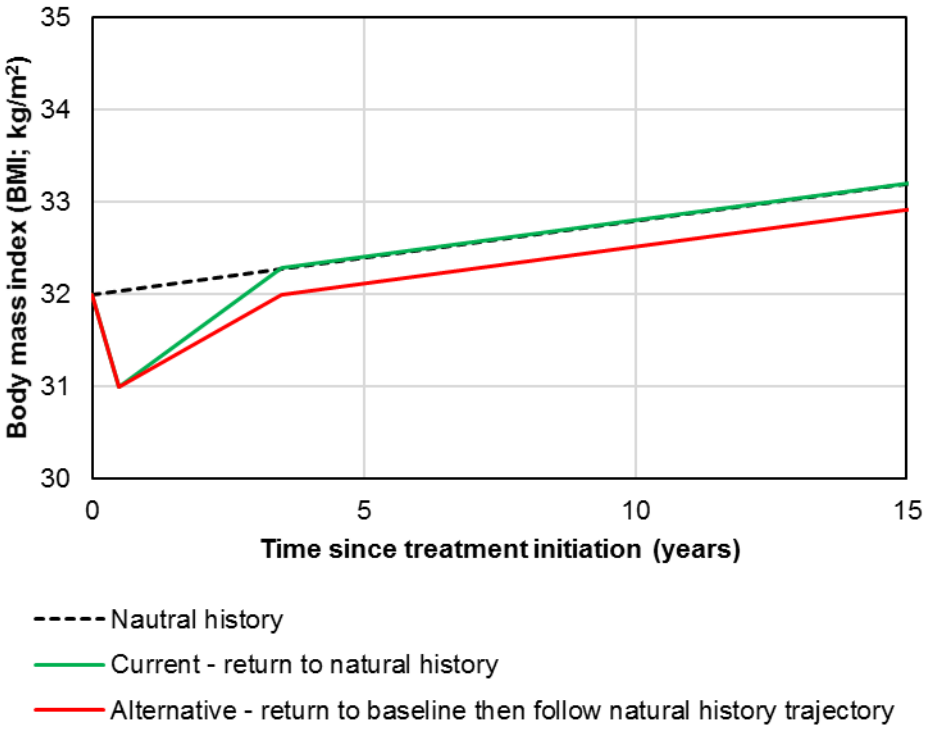
a. Please justify why weight regain towards the predicted BMI (with the natural history model) was preferred over weight regain towards the baseline BMI.

As the ERG are mindful that uncertainty around this parameter has been key in previous analyses, so were we.

Simulated patient weight upon model entry was consistent with the BMI trajectory analysis reported by Ara *et al.*, as described in Section 5.3.1 of the company submission (CS). In the long-term, following treatment discontinuation, for a simulated patient's BMI to be consistent with their characteristics, it was a logical assumption for patients to trend towards their BMI trajectory following discontinuation. As is clear from Figure 9, this was a conservative assumption in comparison to assuming patients reverted to baseline BMI.

In this case, as throughout the model, we made conservative assumptions to prioritise consistency and logic, to illustrate the likely minimum benefit of NB32 adjunct therapy in an area in which it is very challenging to capture and to quantify down-stream health and cost benefits.

Figure 9: BMI projections over time; revert to natural history model versus revert to baseline BMI, with identical gradients from this point



Key: BMI, body mass index.

b. Please justify why the period of three years (for linear weight regain) is appropriate.

Linear regain is one of several alternative assumptions used to inform the process of regain in previous analyses.[7-13] Others have included per-monthly weight regain,[14-17] and assuming a proportion of/all weight loss is maintained indefinitely.[15, 18-20]

The use of a weight regain rate per specified time period (e.g. 0.3kg per month) was considered inappropriate as it does not consider the variability of weight regain by weight

loss – e.g. a patient who lost 3kg would regain weight over a 10-month period, whereas a patient who lost 18kg would regain weight over a 5-year period; the latter of which may be considered relatively long.

Considering a proportion of weight loss maintained (e.g. 20% of weight loss achieved maintained indefinitely) was considered to be inappropriate as there are no data to suggest weight loss achieved through pharmacological weight loss interventions is maintained indefinitely – primarily due to the lack of available long-term follow-up data to provide evidence for this.

Ara *et al.* assumed linear weight regain over 3 years, based on a previous NICE recommendation.[6, 21] This assumption was upheld in four previous studies identified by Ara *et al.,*[8, 10-12] as well as the *de novo* model constructed by Ara *et al.*[6] In the absence of data, consistency across relevant appraisals in considered a valid justification. Exploratory analyses presented in Table 79 in the submission dossier explored the sensitivity of results to the assumed speed of regain; the incremental cost-effectiveness ratio (ICER) versus standard management alone remained below £14,200 when the assumed time to regain was reduced to 2 years.

c. Please justify why weight regain towards the predicted BMI (with the natural history model) was only started after discontinuation of **all treatments** instead of after discontinuation of **active treatments** as assumed by Ara *et al*.[6]

Within the COR trial programme, and reflected in the economic analysis, patients that received placebo (i.e. standard management alone) achieved modest, but evident, weight loss outcomes.

Given that both NB32 and orlistat are provided as an adjunct to standard management, it was important to understand whether NHS patients would, in practice, continue to receive standard management following discontinuation of either NB32 or orlistat.

Email correspondence with Professor John Wilding sought to address the question: *"If* [a patient] *discontinue*[s] *adjunct pharmacotherapy in practice, does* [their] *non-drug therapy ("standard management") also cease?"*. Professor Wilding's response was: *"In practice the standard management would continue…"*.

To illustrate the point further, Professor Wilding made reference to a publication by Sjostrom *et al.* regarding an RCT of orlistat.[22] In this study, patients could switch from orlistat to placebo, and there was continued evidence of weight loss (or "weight loss maintenance"). Professor Wilding also said that *"if lifestyle intervention is stopped then weight regain is also*

common…", though did not have a direct example of this to hand at the time of email correspondence.

Ara *et al.* did not consider non-pharmacological treatment within their scope (the objective of the study was to evaluate the clinical effectiveness and cost-effectiveness of three pharmacological interventions in obese patients[6]), which perhaps meant that this important clinical assumption was not considered. Of course, such assumptions are of more relevance and importance when direct comparisons are required between pharmacological adjuncts to standard management and standard management alone.

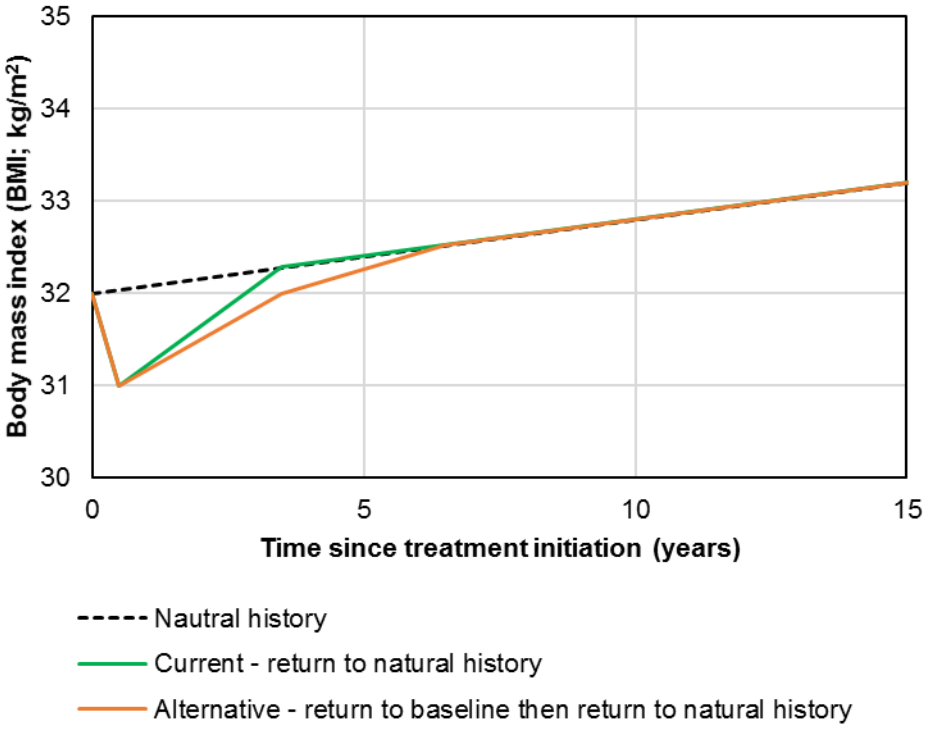
d. Please provide a scenario analysis assuming start of weight regain after discontinuation of **active treatments**.

As discussed in response to Question B1 c, it would be inappropriate to consider modelling weight regain only after the cessation of active pharmacological treatment. If this were applied, it would be implicitly suggested that there is no benefit for patients receiving standard management without pharmacological adjunct (which captures what many NHS England patients currently receive), and that such treatment would lead to no change in "natural history" weight gain.

e. Please provide a scenario analysis, similar to Ara et al.'s[6] base case, in which patients revert to their baseline BMI in three years and then enter the natural history model.

As discussed in response to Question B1 a, the application of weight regain was a conservative assumption in comparison to assuming patients reverted to baseline BMI. An alternative method was explored to demonstrate the robustness of the model to this assumption. In this scenario, patients regained weight to their baseline BMI and re-joined the natural history trajectory upon experience of the next event in the model. This is demonstrated in Figure 10.

Figure 10: BMI projections over time; revert to natural history model versus revert to baseline BMI, then return to identical natural history model after next event



Key: BMI, body mass index.

The results of running the model with this scenario are provided in Table 6.

Table 6: Scenario analysis – patients return to baseline BMI

	Total			Incremental			ICER	
	Costs	**LYs**	**QALYs**	**Costs**	**LYG**	**QALYs**	**vs SM**	**Incremental**
Base case results								
SM	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4148	£294	0.0383	0.0531	£5,538	£5,538
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£13,647	£32,084
Patients return to their baseline BMI following weight regain								
SM	£6,578	33.2638	15.2288					
ORL	£6,868	33.2985	15.2804	£291	0.0348	0.0516	£5,633	£5,633
NB32	£7,618	33.3144	15.3027	£749	0.0159	0.0223	£14,079	£33,620

Key: BMI, body mass index; ICER, incremental cost-effectiveness ratio; LYG, life years gained; ORL, orlistat; QALYs, quality-adjusted life years; SM, standard management.

It should be noted that implementation of this scenario within the model introduces a further conservative assumption. This is because patients receiving standard management alone will discontinue treatment ahead of, or at the same time as, their adjunctively-treated counterparts. As such, the time predicted to their next event following regain will be determined when the patient is relatively younger with a BMI the same as at baseline. As

such, patients treated with standard management alone will generally revert back to the natural history model at a later time than those treated with NB32 or orlistat.

Therefore, the ICERs produced within the scenario analysis are slightly higher than those in the model base case, but are broadly similar, further demonstrating the robustness of the model to assumptions regarding weight regain.

f. Please provide a scenario analysis combining d and e: patients revert to their baseline BMI in three years and then enter the natural history model and start of weight regain after discontinuation of active treatments.

As discussed in response to questions B1c and d, it would be inappropriate to consider modelling weight regain only after the cessation of active pharmacological treatment.

B2. **Priority**: The 16-weeks treatment discontinuation for NB32 was linearly scaled to 12 weeks and assumed to be equivalent to treatment discontinuation used for orlistat.

a. Please justify why the treatment discontinuation for NB32 (linearly scaled or not linearly scaled) is applicable to orlistat.

Treatment discontinuation data for patients receiving NB32 were available from the COR trial programme, as were data for patients receiving standard management alone via the placebo arms of these studies.

As described in the submission dossier, it was essential that we had access to patient-level treatment discontinuation data from these trials to accurately capture treatment costs. These data were analysed to reflect the treatment stopping rules imposed by the EMA that did not originally feature within the COR trial programme studies. To do this, analysis of patient-level data from each of the studies were combined, with the stopping rules retrospectively imposed to estimate the expected TTD for patients in clinical practice. Further information regarding the derivation of TTD for NB32 patients is presented in Section 5.3.2.1 of the manufacturer's submission.

We would have loved to have had similar data available for orlistat patients, but, of course, this was not the case and we were reliant on publicly available data. TTD data were not routinely reported or reported in usable forms in the identified orlistat study publications. Of course, a key limitation in this was that none of the orlistat studies accounted for the EMA treatment stopping rule used for this product.

As such, as the most reasonable assumption possible, it was assumed that patients receiving orlistat would follow a similar trajectory to those receiving NB32. This assumption is inherently conservative given the known toxicity profile of orlistat and its association with treatment discontinuation (via mainly gastrointestinal side effects).

The linear scaling approach was used to account for the different time to primary assessment following treatment initiation (16 weeks for NB32, 12 weeks for orlistat).

It is acknowledged that the assumptions regarding the application of TTD are imperfect, but are not biased in the favour of NB32, and are the most appropriate assumptions to make in the absence of more appropriate data.

b. Please provide a scenario analysis using the NB32 treatment discontinuation for orlistat without linear scaling.

We did consider such an approach during model development, but found its assumptions to be inherently flawed. This analysis would suggest that the discontinuation rate for patients after 1 week of treatment with NB32 are equivalent to the discontinuation rate for patients after 1 week of treatment with orlistat. Given that the first 4 weeks of NB32 are a titration period, imposing such an assumption is theoretically weak.

It is acknowledged that scaling the TTD of NB32 patients to match the 12-week period for orlistat patients does not completely overcome this issue, but such scaling does reduce the proportion of time over which these comparisons are made (i.e. a non-maximum recommended daily dose [RDD] of NB32 is compared with the maximum [RDD] of orlistat for the first 3 weeks).

In addition, it may be expected that due to the difference in safety profiles, a larger proportion of orlistat patients may discontinue treatment ahead of primary assessment versus those treated with NB32. As such, the current active treatment discontinuation assumptions can be considered conservative. Alternative assumptions are possible, but any that would further bias against NB32 would likely not be helpful for decision-making.

c. At the end of the trial follow-up period, it is assumed that all patients would discontinue treatment. Please justify this assumption further and provide a scenario analysis using parametric survival models applied to the COR trial data to extrapolate treatment discontinuation.

Within the *de novo* model, the assumption that treatment discontinuation occurs at the last observable data point is (i) conservative, and (ii) consistent with the approach used within the study by Ara *et al.*

Following successful secondary response assessment, patients are assumed to maintain their weight loss achieved within the first 56 weeks of treatment. No further weight loss is modelled due to the lack of available data to inform the benefits of prolonged treatment. The application of weight loss maintenance is consistent with the minimum requirement for

continued treatment, as discussed with Professor Wilding who stated that treatment would *"be continued in patients who have achieved >5% weight reduction from baseline"*.[5]

In the base case of the study by Ara *et al.*, *"all patients were withdrawn from active treatment at 12 months as this* [was] *the end point for* [their] *evidence."*[6] Therefore, the same approach was considered within our base case model assumption that all active treatment is withdrawn within the observed period of the NB-CVOT study.

As stressed in the submission, the conservative nature of these assumptions are further confounded by the limitations in (i) diseases linked to BMI and (ii) that we only capture delays in time-to and not probability-of events.

The scenario requested suggests to apply parametric survival models to reflect the TTD of patients following secondary assessment. There are several issues with the request to implement such a scenario in the model discussed below.

Professor Wilding noted that *"harnessing evidence from the* [NB-CVOT] *study may be sensible to best inform what happens beyond 12 months."*[5] To introduce assumptions regarding what would happen to patients beyond the duration of follow-up observed in the NB-CVOT data would be guesswork. As such, the current approach within the model was considered in line with clinical expert opinion, while avoiding the imposition of alternative assumptions pertaining to the long-run outcomes.

There is no evidence to support the choice of curve fits regarding long-term extrapolation. Given the large variation in the long-term prediction alternative parameterisations typically yield, the use of any parametric curve was deemed inappropriate as the only criteria from which a decision regarding the "best fitting" curve that could be made would be based on analysis of the statistical goodness of fit (e.g. Akaike or Bayesian Information Criterion).

d. Please clarify what determines whether patients can continue standard management after they have discontinued pharmacological treatment and clarify how time to discontinuation of standard management is subsequently estimated for these patients.

In line with the response to Question B1c, available evidence identified by Professor Wilding states that patients who continue to receive standard management treatment are able to maintain outcomes achieved while receiving adjunctive pharmacological treatment, and therefore *"In practice the standard management would continue…"* following cessation of active pharmacological therapy.[5] Therefore, the model allows patients to continue standard management after they have discontinued pharmacological treatment.

Patients may continue standard management after discontinuation of adjunctive pharmacological therapy (i.e. NB32 or orlistat), dependent on whether they are sampled to

have a longer TTD for adjunctive pharmacological therapy or standard management. Within the model, these are sampled independently, as the likelihood of patients actively continuing standard management following discontinuation of study treatment is varied. To clarify, within the model the time to discontinuation of standard management is sampled upon entry to the model. If the sampled time of discontinuation of adjunctive therapy occurs before this time, the patient will continue to receive standard management treatment. Alternatively, if the sampled time of discontinuation of adjunctive therapy occurs after this time, the patient will continue to receive standard management until discontinuation of adjunctive therapy at which point all treatment will stop.

e. Please justify why it was appropriate to use the NB-CVOT study to estimate treatment discontinuation of standard management beyond 56 weeks despite the difference in population compared with the COR trial programme (which was used for estimating treatment discontinuation of standard management up to 52 weeks). Please also discuss the implications of using a more severe patient population for estimating treatment discontinuation post 56 weeks.

The use of NB-CVOT data beyond 56 weeks was the practical alternative to following Ara *et al.* and assuming treatment continuation and benefit for 1 year only.

Data beyond 1 year of follow-up are scarce, as highlighted by the systematic reviews of evidence, and reinforced through conversation with Professor Wilding: *"harnessing evidence from the* [NB-CVOT] *study may be sensible to best inform what happens beyond 12 months."*[5]

As was described in the submission dossier, and the ERG highlight here, the NB-CVOT study was clearly undertaken in an older cohort of patients with increased comorbidity, and therefore likely poorer prognosis (regarding survival, incidence of cardiovascular events etc.) than patients in the COR trial programme. As such, it is expected that patients in the NB-CVOT study discontinued treatment more rapidly than those in the COR trial programme otherwise would have done.

If the analysis could more fully capture downstream benefits of weight loss for the patient group, this would be masking the benefits associated with continued treatment. As the ability of the analysis to capture downstream effects is so severely limited, the delay in time to discontinuation that TTD data from patients with better prognosis would imply would not accurately translate to health benefits in the analysis.

B3. **Priority**: No re-treatment or alternative treatments after treatment discontinuation are assumed in the model.

a. Please justify the assumption of no re-treatment after treatment discontinuation and provide a scenario analysis incorporating re-treatment with active treatments (i.e. NB32 and/or orlistat) and another scenario analysis incorporating re-treatment with standard management.

Retreatment in clinical practice is plausible, as confirmed by Professor Wilding who stated that *"if patients discontinue treatment, they usually receive no further obesity care until at some point they return for their next attempt at weight loss."*[5]

However, data on (i) what patients would receive as retreatment and (ii) how previous treatment would affect retreatment effectiveness are lacking. Would standard management work equally well for patients who had received NB32 plus standard management and patients who had previously received standard management without pharmacological adjunct?

Answering such questions is beyond the evidence base available for this submission.

b. Please justify the assumption of no alternative treatments after treatment discontinuation and provide a scenario analysis incorporating alternative treatments (e.g. bariatric surgery).

It is acknowledged that due to the difference in the mechanism of action of NB32 and orlistat, indirect retreatment could be plausible in clinical practice (e.g. use of orlistat following NB32, or vice versa).

Data on patients using NB32 having previously received orlistat, or vice versa, are unavailable, but the different mechanisms of actions suggest that treatment effects should be independent, and that the results of this analysis could be used to inform treatment decisions for these patients.

Bariatric surgery is not recommended by NICE in the patient population considered within this appraisal. Namely, NICE PH53 (Weight management: lifestyle services for overweight or obese adults) recommends bariatric surgery for patients with a BMI of over 40kg/m^2.[23]

B4. There is no justification in the company submission for why baseline patient characteristics of patients who receive aspirin and patients who receive anti-hypertensive medication are not varied in the generation of profiles in the CS. In the model, the justification reads that these settings 'are disabled as the risk equations by Ara et al. (2012)[6] cause counter-intuitive results (for example, an increase in BMI causing a decrease in the time to death).' However, the ERG would like to highlight that it seems plausible that an increase in BMI would cause a decrease in the time to death.

a. Please provide clarification and justification for this?

Within the risk equations produced by Ara *et al.* (which were subsequently applied within the *de novo* model used to inform this appraisal), a variety of covariates were included that may impact the predicted time-to-event estimates. These included: the use of aspirin and the use of anti-hypertensive medication, along with several other covariates including age, gender and BMI.

Ceteris paribus an increase in BMI is associated with an decrease in the time to death (as well as other events), as stated within the report by Ara *et al. "Results from the seven BMI risk models showed consistent increases in risk due to an increasing BMI."*[6] Therefore, any confounding variables that were counter to this statement were considered erroneous (i.e. if an increase in BMI did not lead to an increase in risk, then there may be some underlying errors within the statistical model produced).

The analysis undertaken by Ara *et al.* considered patients in the GPRD subject to them being at least 18 years of age and having at least three BMI readings of over 27kg/m^2.[6] In addition, BMI readings were considered in the range of 25-60 kg/m^2.[6]

The statement *"for example, an increase in BMI causing a decrease in the time to death"* relates to the fact that on occasion, when a patient has certain covariates (namely, treatment with aspirin and/or anti-hypertensive medicine), the model can produce results that suggest an increase in BMI leads to an increase in the time to death. To avoid potential errors associated with these counter-intuitive results, these settings were disabled in the model.

b. Please provide a scenario analysis in which these parameters are allowed to vary?

Given our response to part a., a scenario with these settings enabled is not provided. We would suggest that the ERG do not consider the results of such a scenario as plausible given the potential for error in the results.

B5. In the company submission it is stated that 'mean change in body weight estimates determines the proportion of responders and non-responders at secondary response.' However, after the primary assessment, responders and non-responders are assigned a mean change in body weight. Specifically, responders at the first assessment for NB32 are assigned an average weight loss of 9.4%. Hence, these responders at the first assessment automatically also meet the response criterion (i.e. ≥5% weight loss) for the second assessment. In other words, NB32 responders at the first assessment are also automatically responders on the second assessment, if they continue treatment.

a. Please clarify how the proportion of responders and non-responders at the secondary assessment are incorporated in the model.

The proportion of patients who respond at secondary assessment are determined following the random sampling of weight loss at secondary assessment. For primary assessment responders, the mean weight loss at Week 56 is 11.7%, with a standard deviation of 7.2%. For each patient, a random draw is taken from a normal distribution with this mean and SD. Of course, the majority of patients when sampled will achieve a weight loss of above 5%. Using the mean and SD, it can be derived that approximately 17% of responders at primary assessment will no longer respond at secondary assessment, and will therefore discontinue treatment.

b. Please clarify whether for NB32 and orlistat, responders at the first assessment are also automatically responders at the second assessment if they continue treatment. If this is the case, justify this assumption and provide a scenario analysis allowing patients to be identified as non-responders at the second assessment.

As discussed in response to Question B5a, the weight loss achieved at secondary response for each patient is sampled within the model assuming that first-order uncertainty regarding estimated weight loss in normally distributed. For a deterministic model run, the average difference between treatments (derived via the indirect treatment comparison [ITC]) is fixed, but the per-patient weight loss from which the relative effect is applied is varied on a per-patient basis. In probabilistic analysis, the CODA sample from the ITC is applied to account for the second-order uncertainty regarding the estimated relative efficacy of NB32, orlistat and placebo treatment.

Results from a scenario allowing patients to be identified as non-responders at the second assessment is not provided here as this is already applied in the model base case.

B6. **Priority**: Please provide two scenario analyses using data on clinical effectiveness and treatment discontinuation derived from the two ITT populations described in Question A13b from the Clinical Effectiveness section: one based on the ITT with weight regain imputation method; and one based on the ITT with baseline-carried forward analysis.

To clarify, it is anticipated that this request relates to Question A19b. If this is incorrect, please provide further explanation of the details regarding the nature of this request.

As discussed in response to Question A19b, the ITT populations are irrelevant for consideration in the *de novo* model, due to the nature in which weight loss outcomes are derived. Therefore, the requested scenarios have not been performed within the *de novo* economic model and the results of these scenarios are not presented here.

B7. In the company submission model, at diabetic onset (and stroke/MI events), time to primary and secondary assessment is recalculated without subtracting the time at which this event occurred. This appears to delay the time to assessment for those patients that experienced the onset of the respective event before either one of the assessments.

a. Please justify why time to assessment was recalculated in this way?

The equations used to calculate the time to response assessment (both primary and secondary) are fixed. For example, after a patient has experienced the "diabetes onset" event, the equation applied to establish their time to primary response assessment at 16 weeks is shown in Equation 1*.

Equation 1*: Derivation of time to assessment at 16 weeks

$$\boldsymbol{IF(di\_first\_assessment\_occurred\_16w = 1, di\_never, 16 * 7)}$$

In Equation 1*, $di\_first\_assessment\_occurred\_16w$ is an indicator variable (or in discretely integrated condition event (DICE) terminology, a "condition") that determines whether or not primary assessment at 16 weeks has previously occurred. If the assessment has already occurred, this condition will assume a value of 1, and thus the patient is not eligible to experience the "primary response assessment at Week 16" event (denoted using the constant di_never). However, if the patient has not already experienced the "primary response assessment at Week 16" event, the equation will return a value of $16 * 7$ (i.e. 16 weeks in terms of days).

The equations incorporated within the model for assessment at fixed time points (such as at Weeks 12, 16, 52 or 56) are unadjusted. This is applied differently to the equation used to derive the time to clinical events or death, which utilise Equation 1 of the manufacturer's submission to re-calculate event times.

b. If the time to assessment was, in fact, a mistake, please provide results of a corrected analysis?

As described in response to Question B7 a, the application of the time to assessment within the model is correct, with further explanation provided regarding how the time to response assessment is derived within the model. Hence, a scenario with corrected analysis is not required.

Comparators

B8. **Priority**: Please add intense behavioural modification as a comparator in the model and provide cost-effectiveness results? Please use the responses to clarification

question A13c as well as modified resource use and costs data to reflect intense behavioural modification as a comparator in the model.

Intense behavioural modification (often termed BMOD) was not listed as comparator within the scope of this appraisal. BMOD is considered within the NHS as a specialist weight management service, for patients with a BMI in the range of 35–39.9kg/m2, or who are considered "Tier 3" patients.[23] The clinical expert opinion of Professor Wilding confirmed that BMOD (particularly in relation to the COR-BMOD study) is *"similar to the type of non-drug treatment used for Tier 3 patients in NHS practice"*.[5]

In addition, Professor Wilding was asked whether it would be plausible to pool data across all four of the COR trial programme studies (i.e. COR-BMOD with non-BMOD studies). Professor Wilding stated that he would *"expect the effect to be additive"* and considered that *"this is reflected in the study results"*.

As discussed in response to Question B6, it is anticipated that the latter part of this request relates to Question A19b. If this is incorrect, please provide further explanation of the details regarding the nature of this request.

Given that BMOD is not a relevant comparator to this appraisal, the requested analysis has not been performed. BMOD is a non-relevant comparator to the appraisal as it was not included within the scope, is not offered to patients who would otherwise be eligible to receive pharmacological weight loss interventions, and any additional benefit observed through BMOD would be expected to be additive based on clinical expert opinion.

Model structure

B9. It is assumed that only 2 strokes, 2 MIs or 1 stroke and 1 MI can occur (with or without T2DM and patients can develop T2DM after the first event). Please justify that this simplifying assumption is plausible, e.g. that a stroke after 2 MIs does not have any important costs and-or quality of life implications?

The model constructed by Ara *et al.* considered a maximum of two cardiovascular events (namely, myocardial infarction [MI] and stroke). No explicit rationale was provided regarding why only a maximum of two events were possible.

Given the population of obese and overweight patients with comorbidities, limited data are available regarding the incidence of multiple cardiovascular events. As such, the strength of data underlying an alternative to this simplifying assumption would be questionable, and there would be a need for further assumptions.

Only a relatively small proportion of patients would be expected to experience more than two cardiovascular events, and the impact of having experienced these events is unlikely to be accurately reflected within the model due to the limitations in both the analysis of health-

related quality of life (HRQL) (which includes a covariate for MI and stroke, but does not disaggregate these covariates by frequency) and the time to event models (as these include a covariate for cardiovascular event history, but no covariate for repeated history).

The response to this question again highlights the conservative nature of the analysis. Excluding further cardiovascular events from consideration within the model is inherently conservative, as the impact of reducing the frequency/delaying the time to additional events only adds to the benefits expected through treatment with NB32 (as these patients experience the most preferential weight loss outcomes).

B10. General population mortality data are used to inform the probability of death beyond follow-up of 15 years. Please justify this assumption and provide an alternative scenario analysis without this assumption.

Ara *et al*. stated *"For individuals in the event-free health state, the Weibull curves derived from the GPRD are used to predict the time to ACM. These curves are valid for up to a maximum of 15 years, after which standard life tables are used."*[6] The application of time-to-death in the de novo model is consistent with the recommendations Ara *et al*. based on the stated limitations of their own analysis.

The words of Ara *et al.* should be heeded. The time to event (TTE) models are prone to producing inconsistent results beyond 15 years, and so providing the requested scenario would not be helpful.

It is worth noting that if the projections were consistently logical beyond 15 years, it would benefit the case for NB32 for them to be included. Lower BMI should extend time-to-unfavourable health outcomes. Please recognise that we have been forced into inherently conservative assumptions, and would happily present scenarios relaxing such assumptions if they produced consistent results.

Health related quality of life

B11. Utility scores are derived from a Tobit model from PHE.

a. Please clarify that the utility scores obtained from this model have face validity, e.g. by means of provision and discussion of a table with utility scores associated with experiencing the different (combinations of) health events in the model, for an average patient.

The utility scores utilised within the *de novo* model were taken from a published Public Health England analysis of weight loss interventions by Copley *et al.* (2016).[24] A summary of example patients are provided in Table 7. A comparison may be possible utilising data from a publication by Ara and Brazier (2010) of the UK general population, as shown in Table 8.[25]

Table 7: Summary of example patient utilities

Patient characteristics	Tobit model	
	Male	Female
Healthy, 30 years, BMI = 27	0.92	0.90
Healthy, 50 years, BMI = 27	0.88	0.86
Healthy, 70 years, BMI = 27	0.84	0.81
Diabetic, 30 years, BMI = 27	0.87	0.85
Diabetic, 70 years, BMI = 27	0.77	0.74
History of MI, 30 years, BMI = 27	0.85	0.82
History of MI, 70 years, BMI = 27	0.73	0.70
History of stroke, 30 years, BMI = 27	0.83	0.81
History of stroke, 70 years, BMI = 27	0.72	0.69
History of MI and diabetic, 30 years, BMI = 27	0.78	0.75
History of MI and diabetic, 70 years, BMI = 27	0.65	0.61
History of stroke and diabetic, 30 years, BMI = 27	0.76	0.73
History of stroke and diabetic, 70 years, BMI = 27	0.63	0.59
Healthy, 30 years, BMI = 35	0.88	0.86
Healthy, 50 years, BMI = 35	0.84	0.82
Healthy, 70 years, BMI = 35	0.79	0.76
Diabetic, 30 years, BMI = 35	0.83	0.80
Diabetic, 70 years, BMI = 35	0.71	0.68
History of MI, 30 years, BMI = 35	0.80	0.77
History of MI, 70 years, BMI = 35	0.67	0.63
History of stroke, 30 years, BMI = 35	0.78	0.75
History of stroke, 70 years, BMI = 35	0.65	0.62
History of MI and diabetic, 30 years, BMI = 35	0.72	0.69
History of MI and diabetic, 70 years, BMI = 35	0.57	0.54
History of stroke and diabetic, 30 years, BMI = 35	0.70	0.67
History of stroke and diabetic, 70 years, BMI = 35	0.55	0.52
Key: BMI, body mass index; MI, myocardial infarction.		

Table 8: Summary of general population utilities

Age	General population	
	Male	Female
30 years	0.934	0.913
50 years	0.876	0.855
70 years	0.791	0.770

For the "healthy" population (utilising an average BMI of 27kg/m^2), the Tobit model utilities are approximately equivalent to the general population utilities derived by Ara and Brazier. The remainder of the utilities demonstrate face validity in that they are below those of the general population, and demonstrate appropriate levels of disutility associated with the existence of comorbidities.

b. Please justify that the Tobit model was preferred over the OLS regression model (CS Table 60).

Two possible utility models were available for use within the *de novo* model: the Tobit model and the ordinary least squares (OLS) model.

A Tobit model (often termed a "censored [regression] model") is specifically designed to accurately reflect the distribution of data where censoring is known to apply – for example, at a lower or upper bound (or both). For EQ-5D-3L data valued using the UK tariff, utilities are censored above at 1.00 (equivalent to a response of 11111) and below at -0.594 (equivalent to a response of 33333).

The Tobit model aims to estimate the proportion of patients that are located at each of these bounds, and utilises these within the estimation of the overall utility. In short, the Tobit model acknowledges the censoring limits and treats utilities at these limits separately to those in between.

Conversely, an OLS model simply fits a standard regression model to observed EQ-5D data without considering the censoring limits at -0.594 and 1.00. In literature, an OLS model is typically associated with issues relating to the estimation of utility values are the extremes of the observed data.

For example, a study by Longworth *et al.* (2005) demonstrated that when considering a wide range of utility values, and/or utility values that lie close to censoring bounds, an OLS model fails to provide a good fit to observed data, as shown by Figure 11.[26]

Figure 11: Plot of observed values on the OLS model predicted values (taken from Longworth *et al.* [2005][26])

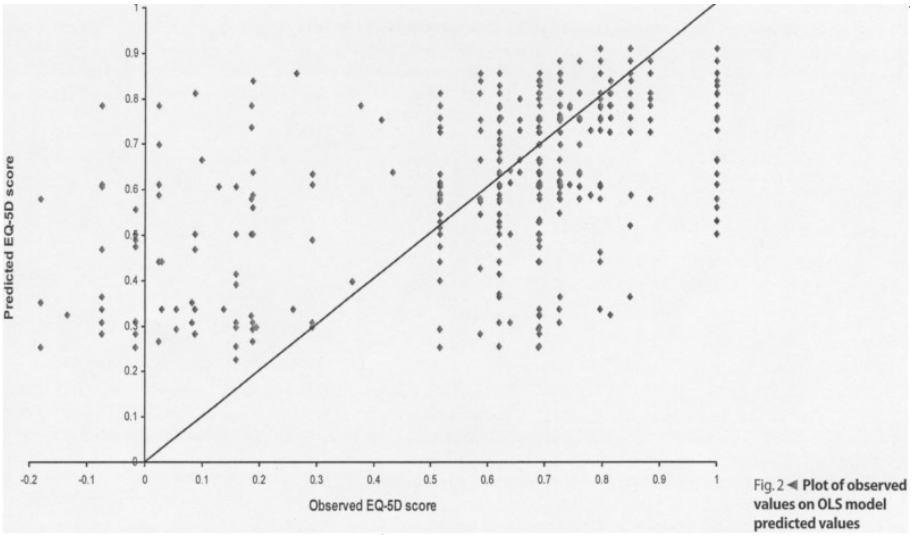


Fig. 2 ◀ Plot of observed values on OLS model predicted values

In a study by Austin *et al*. (2000), *"it was demonstrated that in the presence of a ceiling effect, if the conditional distribution of the measure of health status had uniform variance, then the coefficient estimates from the Tobit model have superior performance compared with estimates from OLS regression. However, if the conditional distribution had non-uniform variance, then the Tobit model performed at least as poorly as the OLS model."*[27] As such, the Tobit model may be considered to perform at least as well as the OLS model.

It is hopefully clear why the Tobit model was selected for the base case. Table 79 of the submission illustrated the effect upon results of alternatively using PHE OLS model results to inform utility assumptions.

B12. Please provide a scenario analysis using the SF-36 data from the COR-II trial.

A scenario utilising SF-36 data from the COR-II study is inappropriate for consideration in decision making for a number of reasons.

Firstly, as reported in the manufacturer's submission, patients completed the SF-36 questionnaire in COR-II only, at baseline, Week 28 and Week 58. Although SF-36 data were collected in COR-II, the frequency of completion and limited follow-up of the COR trials limit the usefulness of these data for the purpose of the economic analysis.

In addition, the SF-36 data available from the COR-II study are inappropriate for modelling utility over a patient's lifetime. Data are only available at three specific time points, and do not reflect the incidence of cardiovascular events (such as MI and stroke) that primarily emerge beyond trail endpoints.

Furthermore, the NICE reference case documents a stated preference for patient-reported EQ-5D data.[28] Copley *et al.* (2016) presented both Tobit and OLS models to predict English patient EQ-5D utility (based on Health Survey for England [HSE] data).[24] Therefore, these utility models were applied within the model given that they were derived using the exact population of relevance to the decision problem.

It is hopefully clear that the requested scenario would not be helpful for decision-making.

B13. Please provide justification for why no utility decrements were applied to adverse events.

Utility decrements were not included within the model due to lack of data available to inform model assumptions. When asked about the expectation of patient HRQL across the three modelled treatment regimens, Professor Wilding stated that he *"would expect patients receiving* [NB32] *to have better on-treatment utility compared to those receiving orlistat, because of its better effectiveness and preferable adverse event profile, and compared to those receiving just diet and exercise intervention* [standard management]*, because of its better effectiveness."*[5]

Therefore, exclusion of disutilities relating to AEs was considered conservative relating to the relative safety profile of NB32 and orlistat. Furthermore, as EQ-5D trial data were unavailable, the positive effects on HRQL of receiving NB32 are not captured within the model. Inclusion of disutility relating to AEs without accounting for the potential benefits in HRQL of receiving an efficacious active treatment (for example, hope of weight loss, increase in confidence etc.) could in fact exacerbate flawed assumptions.

However, a scenario considering a pragmatic application of on-treatment disutility has been provided. As previously discussed, disutility data were not available to inform the impact of AEs on patient HRQL. To account for this, all AEs were assumed to be associated with a utility decrement of 0.05 that persisted for a 1 week duration (in line with the application of AE costs in the model).

Due to poor granularity of available AE data, all AEs reported in the COR I study were utilised within the model. As such, AEs were reported for patients receiving standard management (SM). Therefore, the total weekly utility decrement for NB32 and SM was calculated and the difference was applied to NB32 and orlistat patients while receiving adjunctive therapy.

The results of this scenario are provided in Table 9. As shown, the inclusion of AE-related utility decrements are not associated with having a large impact on HRQL, and hence the overall quality-adjusted life years (QALYs) of patients receiving NB32. The impact of AEs on patients treated with orlistat is expected to be under-represented in the scenario conducted,

due to the difference in toxicity profiles (driven by the mechanisms of action) and lack of comparable safety data.

Table 9: Scenario analysis – inclusion of AE-related utility decrements

	Total			Incremental			ICER	
	Costs	**LYs**	**QALYs**	**Costs**	**LYG**	**QALYs**	**vs SM**	**Incremental**
Base case results								
SM	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4148	£294	0.0383	0.0531	£5,538	£5,538
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£13,647	£32,084
Inclusion of AE-related disutilities for patients treated with NB32 and orlistat								
SM	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4139	£294	0.0383	0.0523	£5,625	£5,625
NB32	£7,563	33.5343	15.4372	£750	0.0192	0.0232	£13,820	£32,272

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; ORL, orlistat; QALYs, quality-adjusted life years; SM, standard management.

Resource use and costs

B14. In the company submission, the cost of Diabetes Mellitus are not derived from Ara et al.[6] Please justify this.

The cost of treating T2DM in the study conducted by Ara *et al.* was only applied for the first year (i.e. no follow up costs applied for T2DM beyond a year). This can be seen in Table 25 of the report by Ara *et al.*[6]

It was considered a sensible improvement upon Ara *et al.* to assume downstream T2DM costs. However, it seemed inappropriate to consider applying Ara *et al.* cost every year, since the cost was presented within this report as the cost for the first year only. Therefore, an alternative and more recent annual cost estimate was identified and applied in the model, as described in Section 5.5 of the submission dossier.

B15. Please justify why drug wastage of NB32 was not incorporated in the model.

Most patients in clinical practice are expected to discontinue treatment with NB32 at assessment visits (i.e. at Week 16 and Week 56). The publication by Fujioka *et al.* shows that over the COR trial programme studies, 51% of all randomised patients achieved at least 5% weight loss from baseline at Week 16.[29] At Week 56, a proportion of patients will discontinue treatment, which based on the results of the *de novo* model is expected to be in the region of 17% of patients still on treatment at this time.

As such, the majority of discontinuations are expected to occur at these response assessment visits. Given this, and that discontinuations driven by health professionals will not involve wastage, inclusion of wastage is not anticipated to impact cost-effectiveness results greatly.

To illustrate that the model is robust to this assumption, a scenario considering drug wastage has been performed, with results provided in Table 10. As expected, inclusion of drug wastage does not result in large changes to the model results as the majority of discontinuation occur at response assessment points.

Table 10: Scenario analysis – drug wastage

	Total			Incremental			ICER	
	Costs	**LYs**	**QALYs**	**Costs**	**LYG**	**QALYs**	**vs SM**	**Incremental**
Base case results								
SM	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4148	£294	0.0383	0.0531	£5,538	£5,538
NB32	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£13,647	£32,084
Inclusion of drug wastage								
SM	£6,519	33.4768	15.3616					
ORL	£6,841	33.5151	15.4148	£321	0.0383	0.0531	£6,047	£6,047
NB32	£7,670	33.5343	15.4381	£830	0.0192	0.0234	£15,047	£35,510

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; ORL, orlistat; QALYs, quality-adjusted life years; SM, standard management.

Probabilistic sensitivity analysis

B16. Please provide a justification as to why the company believes that probabilistic sensitivity analysis using 100 simulations results in stable / plausible results.

In order to fully demonstrate the process of determining the number of probabilistic sensitivity analysis (PSA) simulations required, it is necessary to re-iterate the process of model development that lead to the approach utilised within this appraisal.

As discussed in Sections 5.1 and 5.2.2 of the submission dossier, an individual-level model was developed for this appraisal, as an individual-level approach is better suited than a

cohort-level approach to capture the chronic implications of both weight and weight-related health events in a heterogenous group of overweight and obese patients.

The choice of software package was driven by the need to ensure that the model is applicable across various health technology assessment agencies internationally – some of which only consider models constructed within Microsoft Excel®. Microsoft Excel is also the preferred software package of NICE, and is typically considered more transparent than simulation models constructed in other software packages (such as R or Simul8) due to reviewer familiarity with Microsoft Excel.

In consideration of the limitations of requiring both (i) an individual-level model, and (ii) a model constructed in Microsoft Excel, a DICE model was selected as a transparent approach to accurately demonstrate the costs and effects associated with NB32 treatment. The model constructed in Microsoft Excel is limited in regards to processing power to simulate very large cohorts of patients.

PSA considers the joint uncertainty of all model parameters by sampling individual parameters from their respective distributions. As highlighted in the manufacturer's submission, to establish how many patient profiles are required to produce stable model results, a diagnostic exercise was carried out that yielded a minimum of 500 patients before results generally stabilised.

Regarding the number of PSA simulations, a trade-off between the number of patient profiles and the number of probabilistic draws was made. Given that within the PSA, 500 patient profiles are used for each PSA run, the number of PSA runs was chosen at 100. This number of PSA iterations was associated with a long run-time due to the limitations in processing power associated with Microsoft Excel.

It is important to acknowledge that much of the key uncertainty regarding model results is structural and methodological, and based on the key conservative assumptions underpinning the analysis. The uncertainty regarding results stemming from such uncertainty is not illustrated by PSA or deterministic sensitivity analyses. Therefore, the choice of the number of PSA iterations (be that 100, 100,000 or 100,000,000) does not demonstrate how conservative the structural model assumptions are.

Not all model parameters were able to be considered within the PSA – namely, the uncertainty surrounding the key time to event equations reported by Ara *et al.* were unable to be considered as variance-covariance matrices for these equations was not provided in time for the submission following an email request for these. It was appreciated that staff leave and movement of key staff since Ara *et al.* publication may have been a factor in obtaining these matrices.

In summary, 100 simulations is likely too few runs to directly establish the extent of parameter uncertainty within the model. Importantly, the parameter uncertainty of the key TTE risk equations is not explored via this analysis, and the majority of the key uncertainty is structural and methodological, as opposed to directly related to model parameters, and therefore PSA is inappropriate to describe the majority of the uncertainty expected within the model.

Cost effectiveness results

B17. Please provide an overview of the disaggregated costs, QALYs and LYs (using the conditions specified in company submission Figure 25).

As discussed in Section 5.7.3, disaggregated results for life years (LYs) and QALYs are not available from the model. This is a direct consequence of the chosen model structure, as there are no distinct health states for which disaggregated LYs and QALYs can be summarised. Disaggregated costs by "health state" are also not possible, but are possible to present according to cost item (e.g. drug costs, AEs etc.) that are presented in the manufacturer's submission in Table 76.

Figure 25 of the manufacturer's submission provides an overview of the model structure. To clarify, the structure demonstrates the possible events that patients may experience, and does not demonstrate the health states patients may enter and exit. As the model considers events, the time between events is not reflective of a specific health state (as the ordering of events can vary for each patient).

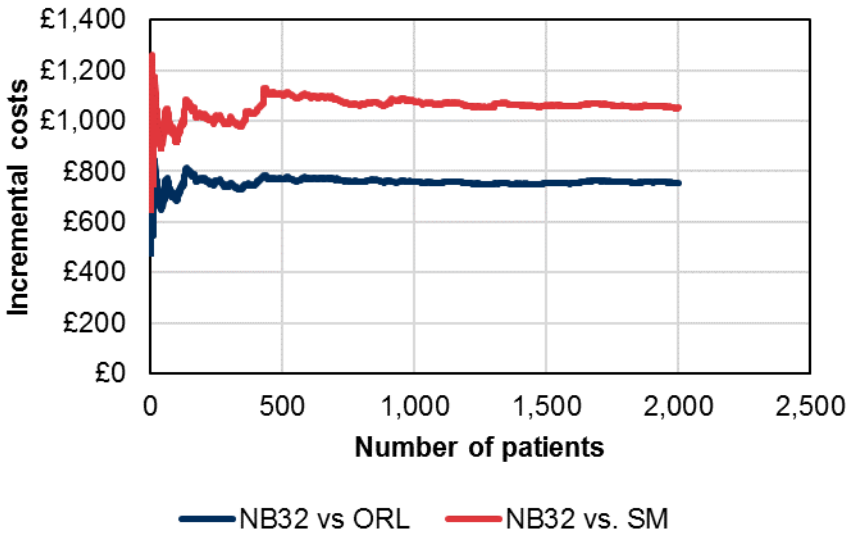
In short, disaggregated results according to the structure presented in the model schematic are not possible to produce due to the structure of the event-based model; hence, are not provided in response to this question.

B18. Ara et al.[6] used a cohort of 1,000,000 patients in their patient-level simulation and stated that, with a cohort size of 200,000 patients, there was still a small amount of variation in results, which stabilised after simulation of 400,000 patients. In contrast, a cohort of only 1,000 patients was used in the company submission. Company submission Figures 34 and 35 provide a diagnostic exercise to examine the minimum number of patients needed to obtain stable results.

   a. Please provide similar figures using the incremental costs, incremental QALYs and the ICER (QALYs) and justify why 1,000 patients were deemed sufficient.

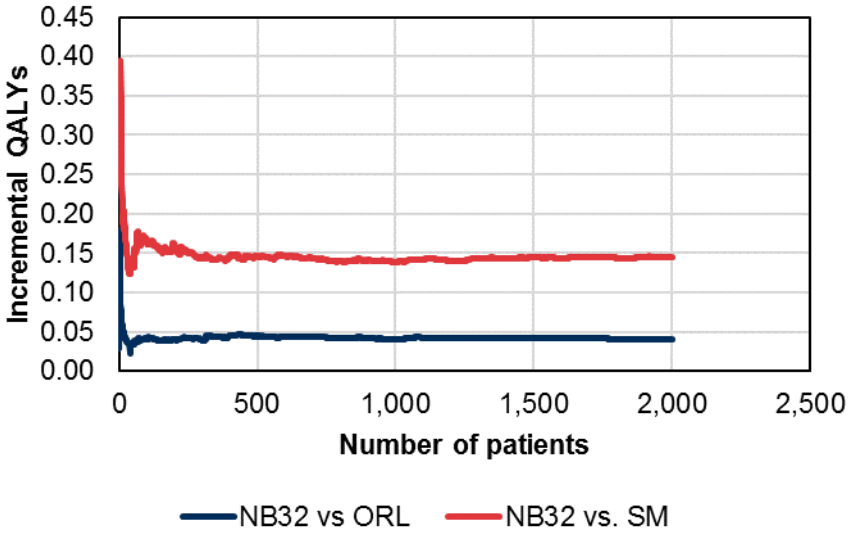
Below are the requested plots using the incremental costs, incremental QALYs and the ICER (QALYS), shown in Figure 12, Figure 13 and Figure 14, respectively.

Figure 12: Diagnostic exercise – incremental costs



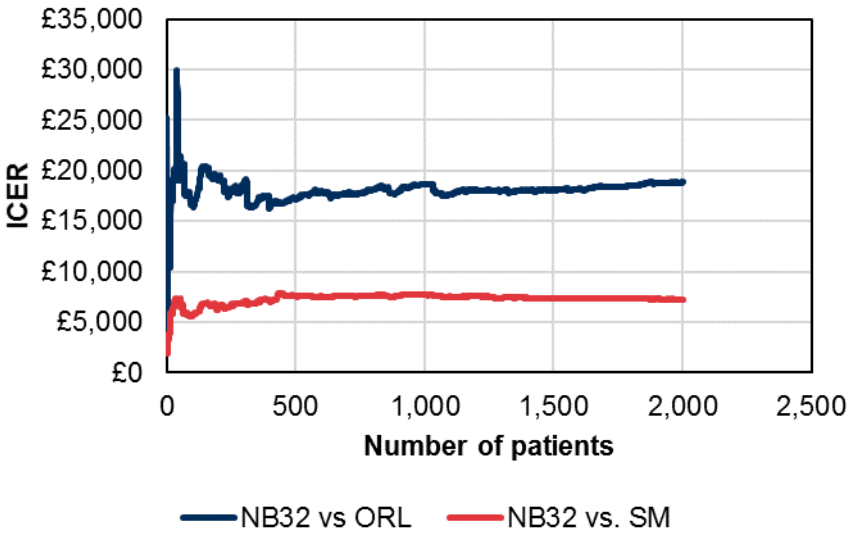
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

Figure 13: Diagnostic exercise – incremental QALYs



Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

Figure 14: Diagnostic exercise – ICER



Key: ICER, incremental cost-effectiveness ratio; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

By considering the incremental plots (shown in Figure 12 and Figure 13 for costs and QALYs, respectively), it can be seen that beyond approximately 500 patients, the incremental costs and QALYs stabilise. The incremental plots do not demonstrate the uncertainty in the total costs and QALYs per treatment arm (which are presented in the manufacturer's submission in Figure 34 and Figure 35, respectively). The total costs and QALYs plots demonstrate further degrees of variability that appear to stabilise slightly later at approximately 1,000 patients.

The final plot for the ICER (shown in Figure 14) should be interpreted with caution. The joint uncertainty in both the incremental costs and incremental QALYs can appear to over-inflate the overall uncertainty in model results, as the ratio of both quantities can appear to present large uncertainty in model results; whereas, in reality, this is a caveat of the sensitivity of the ICER ratio, particularly in the presence of relatively small incremental QALY gains.

However, the ICER plot does demonstrate that results stabilise after approximately 500 patients for the comparison of NB32 and SM. For the comparison of NB32 and orlistat, the results are less certain, but do stabilise after approximately 1,000 patients.

b. Please justify the usage of 1,000 patients given that Ara et al[6] used a cohort of 1,000,000 patients and stated that, with a cohort size of 200,000 patients, there was still a small amount of variation in results.

The model constructed by Ara *et al.* was a cohort simulation model developed in Simul8 version 17.0 build 2277.[6] As discussed in response to Question B16, the model constructed

to inform this submission was constructed in Microsoft Excel. Therefore, the model constructed to inform this appraisal was limited in regards to processing power and run time. The model by Ara *et al.* is able to simulate a very large cohort of patients given the specialist software, which does not require patient characteristics to be controlled for within each run. However, we were able to avoid the need to produce model results for a very large cohort (such as 1,000,000 patients) by controlling baseline characteristics for each model run. By controlling these characteristics, the only difference across patients was the treatment received.

As such, though comparison of the patient numbers considered across studies is startling at face value, the models do not consider the sampling of patients in the same way, and therefore should not be compared as like-for-like.

Validity

B19. Please provide the results of the internal validation described at the end of company submission section 5.10.

The internal validation included, but was not limited to, a checklist of basic validity checks (e.g. setting all costs to zero and ensuring the model outputs zero costs), sheet by sheet check of model logic (e.g. checking DICE equation logic), module by module check of VBA logic, validity assessment of outcomes (e.g. comparing available trial data with the outcomes of the model), and editorial checks (e.g. performing a spell check of model content).

The checklist itself is commercial property, and therefore we are unable to provide the completed checklist.

B20. Please provide the source for and justify the validity of equation 1 in the company submission. Additionally, provide a simple example using this formula and explain why the results are plausible.

An explicit reference is not available for the equation, as it is simply a weighted average of predicted times. The equation accounts for what proportion of time is yet to be observed (i.e. time between originally calculated TTE and current time), and uses this to weight the additional expected time to an event based on updated risk factors.

An example of this in practice is provided below:

Consider the situation where at baseline, a patient is predicted to die in 20 years. At one year, the patient has experienced a change in risk factor(s) (e.g. weight loss) such that their predicted time to death is expected to increase (to say 21 years). Therefore, we may have:

$$Time_{current} = 1\ year;\ TTE_{original} = 20\ years;\ TTE_{re-sample} = 21\ years$$

We need to adjust the time to death from baseline considering that at 1 year, the patient experiences a change in their risk factor(s) such that their time to death changes (increases). Using the formula in Equation 1 we obtain:

$$TTE_{recalculated} = Time_{current} + TTE_{re-sample}\left(\frac{TTE_{original} - Time_{current}}{TTE_{original}}\right)$$

$$TTE_{recalculated} = 1 + 21\left(\frac{20-1}{20}\right) = 20.95\ years$$

This formula accounts for the following information:

- *The patient has survived up until 1 year*
- *The patient was originally predicted to survive until 20 years*
- *The patient is predicted to survive until 21 years, but has already "achieved" 1 year of this survival from the equation using previous risk factors.*

This formula avoids the following possible erroneous applications of time to event models:

- *Using the re-sampled time to death from the current time point, as this suggests the patient dies at time = 22 years.*
- *Using the original time to death (20 years), as this does not take into account the change in the predicted time to event.*
- *Using the revised time to death (21 years), as this does not take into account the previous period in which the equation is invalid (e.g. patient does not immediately lose weight).*

Therefore, adjusting the original time to event by re-sampling and considering the current point in time allows estimation of the overall time to event accounting for the change in risk factors (such as weight loss, for example).

It should be noted that the equation does not apply to every event within the model, as some events occur at fixed time points. For example, primary assessment for patients treated with orlistat is always at exactly 12 weeks after treatment initiation.

Section C: Textual clarifications and additional points

None

References

1. Orexigen. A Multicenter, Randomized, Double Blind, Placebo Controlled Study Comparing the Safety and Efficacy of Two Doses of Naltrexone Sustained Release

(SR)/Bupropion Sustained Release (SR) and Placebo in Obese Subjects. (Clnical Study Report) 28 January 2010.

2. Orexigen. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Obese Subjects. (Clnical Study Report) 2010.

3. Orexigen. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone Sustained Release (SR)/Bupropion Sustained Release (SR) and Placebo in Subjects with Obesity Participating in a Behavior Modification Program. (Clinical Study Report) 4 January 2010.

4. Orexigen. A Multicenter, Randomized, Double-Blind, Placebo-Controlled Study Comparing the Safety and Efficacy of Naltrexone 32 mg Sustained Release/Bupropion 360 mg Sustained Release and Placebo in Obese Subjects with Type 2 Diabetes Mellitus. (Clinical Study Report) 2009.

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

6. 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.

7. Hampp C, Hartzema AG and Kauf TL. Cost-utility analysis of rimonabant in the treatment of obesity. *Value Health*. 2008; 11(3):389-99.

8. Lacey LA, Wolf A, O'Shea D, et al. Cost-effectiveness of orlistat for the treatment of overweight and obese patients in Ireland. *Int J Obes (Lond)*. 2005; 29(8):975-82.

9. Jaime Caro J, Ozer Stillman I, Danel A, et al. Cost effectiveness of rimonabant use in patients at increased cardiometabolic risk: estimates from a Markov model. *Journal of Medical Economics*. 2007; 10(3):239-54.

10. Hertzman P. The cost effectiveness of orlistat in a 1-year weight-management programme for treating overweight and obese patients in Sweden : a treatment responder approach. *PharmacoEconomics*. 2005; 23(10):1007-20.

11. Ruof J, Golay A, Berne C, et al. Orlistat in responding obese type 2 diabetic patients: meta-analysis findings and cost-effectiveness as rationales for reimbursement in Sweden and Switzerland. *Int J Obes (Lond)*. 2005; 29(5):517-23.

12. Maetzel A, Ruof J, Covington M and Wolf A. Economic evaluation of orlistat in overweight and obese patients with type 2 diabetes mellitus. *PharmacoEconomics*. 2003; 21(7):501-12.

13. Oster G, Thompson D, Edelsberg J, et al. Lifetime health and economic benefits of weight loss among obese persons. *Am J Public Health*. 1999; 89(10):1536-42.

14. Ara R and Brennan A. The cost-effectiveness of sibutramine in non-diabetic obese patients: evidence from four Western countries. *Obes Rev*. 2007; 8(4):363-71.

15. Brennan A, Ara R, Sterz R, et al. Assessment of clinical and economic benefits of weight management with sibutramine in general practice in Germany. *Eur J Health Econ*. 2006; 7(4):276-84.

16. Warren E, Brennan A and Akehurst R. Cost-effectiveness of sibutramine in the treatment of obesity. *Med Decis Making*. 2004; 24(1):9-19.

17. James WP, Astrup A, Finer N, et al. Effect of sibutramine on weight maintenance after weight loss: a randomised trial. STORM Study Group. Sibutramine Trial of Obesity Reduction and Maintenance. *Lancet*. 2000; 356(9248):2119-25.

18. Burch J, McKenna C, Palmer S, et al. Rimonabant for the treatment of overweight and obese people. *Health Technol Assess*. 2009; 13 Suppl 3:13-22.

19. van Baal PHM, Hoeymans N, Hoogenveen RT, et al. Disability weights for comorbidity and their influence on Health-adjusted Life Expectancy. *Population Health Metrics*. 2006; 4:1-.

20. Roux L, Kuntz KM, Donaldson C and Goldie SJ. Economic evaluation of weight loss interventions in overweight and obese women. *Obesity (Silver Spring)*. 2006; 14(6):1093-106.

21. O'Meara S, Riemsma R, Shirran L, et al. A rapid and systematic review of the clinical effectiveness and cost-effectiveness of orlistat in the management of obesity. *Health Technol Assess*. 2001; 5(18):81.

22. Sjostrom L, Rissanen A, Andersen T, et al. Randomised placebo-controlled trial of orlistat for weight loss and prevention of weight regain in obese patients. European Multicentre Orlistat Study Group. *Lancet*. 1998; 352(9123):167-72.

23. National Institute of Health and Care Excellence (NICE). PH53: Weight management: lifestyle services for overweight or obese adults. 2014. Available at: https://www.nice.org.uk/guidance/ph53. Accessed: 12 October 2016.

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

25. Ara R and Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health*. 2011; 14(4):539-45.

26. Longworth L, Buxton MJ, Sculpher M and Smith DH. Estimating utility data from clinical indicators for patients with stable angina. *Eur J Health Econ*. 2005; 6(4):347-53.

27. Austin PC, Escobar M and Kopec JA. The Use of the Tobit Model for Analyzing Measures of Health Status. *Quality of Life Research*. 2000; 9(8):901-10.

28. National Institute of Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. 2013. Available at: https://www.nice.org.uk/process/pmg9/chapter/foreword. Accessed: 1 September 2016.

29. Fujioka K, Plodkowski R, O'Neil PM, et al. The relationship between early weight loss and weight loss at 1 year with naltrexone ER/bupropion ER combination therapy. *Int J Obes (Lond)*. 2016; 40(9):1369-75.

Single Technology Appraisal

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

Dear Liv,

Please find enclosed the further analysis with respect to change in BMI as additional clarification to the Evidence Review Group, Kleijnen systematic reviews, questions.

Please let me know if you have any additional questions.

Yours sincerely

Hans-Joerg Fugel

Table 1: Change in BMI from baseline for trial NB-301 (COR-1)

Adhoc 878: BMI (kg/m2), Change from Baseline to Endpoint (LOCF)
Full Analysis Set
Double-Blind Treatment Phase
NB-301

		Baseline					Endpoint					Change				
Treatment	n	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max
1) Placebo	511	36.18	4.03	27.50	35.80	45.70	35.73	4.53	22.40	35.30	49.70	-0.45	2.09	-10.20	-0.30	10.50
2) NB16	471	36.27	4.23	27.40	35.80	46.80	34.48	4.75	21.10	34.20	48.60	-1.79	2.49	-14.40	-1.20	5.00
3) NB32	471	36.17	4.41	27.20	35.80	45.90	33.99	5.04	23.00	33.50	47.70	-2.18	2.58	-12.60	-1.70	6.30

Main Effects (Type III Sums of Squares)	Raw Data		
Treatment	F = 69.91	df = 2	p = <.001
Pooled Study Center	F = 0.94	df = 31	p = 0.567

Least Squares for Change from Baseline		
1) Placebo	-0.48	(SE=0.11)
2) NB16	-1.80	(SE=0.11)
3) NB32	-2.20	(SE=0.11)

Pairwise Comparison of LS Means				
NB16-Placebo	Difference = -1.33	Two-sided 95% CI: (-1.63, -1.03)	t = -8.65	p = <.001
NB32-Placebo	Difference = -1.72	Two-sided 95% CI: (-2.02, -1.42)	t = -11.23	p = <.001

Table 2: Change in BMI from baseline for trial NB-302 (COR-BMOD)

TOS2010-A-302-BMI-1: BMI (kg/m^2) Change from Baseline to Endpoint (LOCF)
Full Analysis Set
NB-302

		Baseline					Endpoint					Change				
Treatment	n	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max
1) Placebo	193	36.95	4.21	28.70	36.80	45.50	34.95	5.22	22.60	34.30	46.30	-2.00	2.93	-14.40	-1.40	4.60
2) NB32	482	36.41	4.14	27.50	35.90	46.00	32.90	5.06	21.40	32.50	47.70	-3.52	3.09	-18.20	-3.00	5.30

Main Effects (Type III Sums of Squares)

Raw Data

Study Center	F= 4.90	df= 8	p= <.001
Treatment	F= 36.76	df= 1	p= <.001

Least Squares for Change from Baseline

1) Placebo	-1.81	(SE= 0.22)
2) NB32	-3.36	(SE= 0.15)

Pairwise Comparison of LS Means

NB32-Placebo Difference = -1.54	Two-sided 95% CI: -2.04, -1.04	t= -6.06	p= <.001

Table 3: Change in BMI from baseline for trial NB-303 (COR-2)

TOS2010-A-303-BMI-1: BMI (kg/m^2) Change from Baseline to Week 56 Endpoint (Weighted LOCF NB32 vs. Placebo)
Full Analysis Set
NB-303

	Baseline					Endpoint					Change				
Treatment	n	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max
1) Placebo	456	36.15	4.29	27.00	35.60	45.40	35.63	4.54	24.80	34.80	48.00	-0.52	1.78	-9.50	-0.30	5.20
2) NB32	702	36.15	4.37	26.60	35.50	45.50	33.51	5.11	20.80	32.80	48.50	-2.64	2.64	-16.00	-2.10	6.50

Main Effects (Type III Sums of Squares)

		Raw Data	
Treatment	F= 156.56	df= 1	p= <.001
Pooled Study Center	F= 1.87	df= 34	p= 0.002

Least Squares for Change from Baseline		
1) Placebo	-0.46	(SE= 0.12)
2) NB32	-2.26	(SE= 0.09)

Pairwise Comparison of LS Means

NB32-Placebo Difference = -1.80	Two-sided 95% CI: -2.09, -1.52	t= -12.51	p= <.001

Table 4: Change in BMI from baseline for trial NB-304 (COR-DM)

Adhoc 877: BMI (kg/m2), Change from Baseline to Endpoint (LOCF)
Full Analysis Set
Double-Blind Treatment Phase
NB-304

		Baseline					Endpoint					Change				
Treatment	n	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max	Mean	SD	Min	Median	Max
1) Placebo	159	36.37	4.50	26.80	36.50	46.10	35.67	4.53	24.70	35.40	44.90	-0.69	1.70	-7.90	-0.40	3.00
2) NB32	265	36.71	4.75	27.10	36.20	46.30	34.83	5.06	23.80	34.50	50.80	-1.88	2.15	-9.90	-1.60	5.50

Main Effect (Type III Sums of Squares)

Treatment	F = 33.70	df = 1	p = <.001
HbAlc Strata	F = 1.21	df = 1	p = 0.272
Sulfonylurea Strata	F = 0.98	df = 1	p = 0.323

Least Squares for Change from Baseline

1) Placebo	-0.68	(SE=0.16)
2) NB32	-1.84	(SE=0.13)

Pairwise Comparison of LSMeans

NB32-Placebo Difference = -1.16 Two-sided 95% CI: (-1.55, -0.77) t = -5.81 p = <.001

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Patient/carer organisation submission (STA)

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

Thank you for agreeing to give us your views on this treatment that is being appraised by NICE and how it could be used in the NHS. Patients, carers and patient organisations can provide a unique perspective on conditions and their treatment that is not typically available from other sources. We are interested in hearing about:

- the experience of having the condition or caring for someone with the condition
- the experience of receiving NHS care for the condition
- the experience of having specific treatments for the condition
- the outcomes of treatment that are important to patients or carers (which might differ from those measured in clinical studies, and including health-related quality of life)
- the acceptability of different treatments and how they are given
- expectations about the risks and benefits of the treatment.

To help you give your views, we have provided a questionnaire. You do not have to answer every question — the questions are there as prompts to guide you. The length of your response should not normally exceed 10 pages.

1. About you and your organisation

Your name: [redacted]

Name of your organisation: HOOP (Helping Overcome Obesity Problems)

Your position in the organisation: [redacted]

Brief description of the organisation: HOOP (Helping Overcome Obesity Problems) is an obesity Charity, with over 15000 members currently. We are a unique group of passionate people, including professionals, with the common aim to make the changes needed so that all children and adults struggling with obesity are given access to the services which are right for them.

HOOP UK is run by people who care and understand:

Our aim is to help children and adults achieve a happy and healthy life style

We aim to create the correct environment for all to succeed and we aim to offer support every step of the way.

(For example: who funds the organisation? How many members does the organisation have?)

We are asking for your collective view as an organisation and will be asking patient experts for their individual input separately. If you have the condition, or care for someone with the condition, you may wish to complete a patient expert questionnaire to give your individual views as well.

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: None

2. Living with the condition

What is it like to live with the condition or what do carers experience when caring for someone with the condition?

I have been obese since my late teens/early 20's and morbidly obese since my mid 20's. Life is a struggle; there are many restrictions I face in life from needing an extension belt on an aeroplane, to not being able to partake in certain activities such as sky diving, flying in a helicopter etc…, to being judged on my appearance and from people feeling like it is acceptable to give

you their opinion on your well being whether you want it or not! I am also wary of the type of chair I sit on, will it hold me? Or will I fit in it? Where I sit in a restaurant, can I get to the toilet easily without having to squeeze past people or tables?

I have had experience of losing weight, putting it back on again, losing it etc…..it feels like a vicious circle and an endless battle. I can't remember a time in my life where I have been happy with my body and have spent most of my life hating myself. This is just a flavour of life as an obese person….

3. Current practice in treating the condition

Which treatment outcomes are important to patients or carers? (That is, what would patients or carers like treatment to achieve?) Which of these are most important? If possible, please explain why.

What is your organisation's experience of currently available NHS care and of specific treatments for the condition? How acceptable are these treatments and which are preferred and why?

We have found that support from the NHS varies in different areas, some areas offer large amounts of support through tier 3 & 4 services, where others have very little support and in fact penalise you for being over weight by restricting your access to surgery etc….

Services tend to focus on the diet and exercise element of weight loss and very few services focus on a patient's mental health or address the reason why they may be obese. This is the area that needs addressing and until this is, the success of diet and exercise alone is pretty limited.

4. What do patients or carers consider to be the advantages of the treatment being appraised?

Benefits of a treatment might include its effect on:

- the course and/or outcome of the condition
- physical symptoms
- pain
- level of disability

- mental health
- quality of life (such as lifestyle and work)
- other people (for example, family, friends and employers)
- ease of use (for example, tablets rather than injection)
- where the treatment has to be used (for example, at home rather than in hospital)
- any other issues not listed above

Please list the benefits that patients or carers expect to gain from using the treatment being appraised.

It is another option for people to try, currently there is only one approved medication for obesity in the NHS and this is not the most effective or suitable for everyone. There is a real gap in the NHS for the treatment of Obesity, patients are either given diet and exercise advice or referred for bariatric surgery, there doesn't seem to be a middle ground. I really feel like there is a large group of people that may benefit from a medication to help tackle their obesity. By losing just 10% body weight can have a significant impact on a patient's life.

Please explain any advantages that patients or carers think this treatment has over other NHS treatments in England.

There are currently no treatments on the market that address appetite or satiety, so I definitely see a place for such a product.

If you know of any differences in opinion between patients or carers about the benefits of the treatment being appraised, please tell us about them.

5. What do patients and/or carers consider to be the disadvantages of the treatment being appraised?

Disadvantages of a treatment might include:

- aspects of the condition that the treatment cannot help with or might make worse
- difficulties in taking or using the treatment (for example, injection rather than tablets)
- side effects (for example, type or number of problems, how often, for how long, how severe. Please describe which side effects patients might

be willing to accept or tolerate and which would be difficult to accept or tolerate)

- where the treatment has to be used (for example, in hospital rather than at home)
- impact on others (for example, family, friends and employers)
- financial impact on the patient and/or their family (for example, the cost of travel to hospital or paying a carer)
- any other issues not listed above

Please list any concerns patients or carers have about current NHS treatments in England.

Obesity treatments are treating the symptoms, but not addressing the cause. I think a double pronged attack of treatment and some psychological support could be advantageous.

Please list any concerns patients or carers have about the treatment being appraised.

It is still not addressing the head and why someone may overeat.

If you know of any differences in opinion between patients or carers about the disadvantages of the treatment being appraised, please tell us about them.

6. Patient population

Are there any groups of patients who might benefit more from the treatment than others? If so, please describe them and explain why.

Patients that have binge eating disorders.

Are there any groups of patients who might benefit less from the treatment than others? If so, please describe them and explain why.

7. Research evidence on patient or carer views of the treatment

Is your organisation familiar with the published research literature for the treatment?

☐ Yes ☒ No

If you answered 'no', please skip the rest of section 7 and move on to

section 8.

Please comment on whether patients' experience of using the treatment as part of their routine NHS care reflects the experiences of patients in the clinical trials.

Do you think the clinical trials have captured outcomes that are important to patients? Are you aware of any limitations in how the treatment has been assessed in clinical trials?

If the treatment being appraised is already available in the NHS, are there any side effects that were not apparent in the clinical trials but have emerged during routine NHS care?

Are you aware of any relevant research on patient or carer views of the condition or existing treatments (for example, qualitative studies, surveys and polls)?

☐ Yes ☐ No

If yes, please provide references to the relevant studies.

8. Equality

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Protected characteristics are: age; being or becoming a transsexual person; being married or in a civil partnership; being pregnant or having a child; disability; race including colour, nationality, ethnic or national origin; religion, belief or lack of religion/belief; sex; sexual orientation.

Please let us know if you think that recommendations from this appraisal could have an adverse impact on any particular groups of people, such as:

- excluding from full consideration any people protected by the equality legislation who fall within the patient population for which the treatment is/will be licensed;
- having a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the treatment;
- any adverse impact on people with a particular disability or disabilities.

Please let us know if you think that there are any potential equality issues that should be considered in this appraisal.

Access to the product should be the same wherever you live in the country, the NHS have a bit of a postcode lottery going on sometimes and this isn't equitable.

Are there groups of patients who would have difficulties using the treatment or currently available treatments? Please tell us what evidence you think would help the Committee to identify and consider such impacts.

9. Other issues

Do you consider the treatment to be innovative?

X Yes ☐ No

If yes, please explain what makes it significantly different from other treatments for the condition.

There is currently nothing in the market that addresses appetite or satiety, so I see a place for a product like this.

Are there any other issues that you would like the Appraisal Committee to consider?

No

10. Key messages

In no more than 5 bullet points, please summarise the key messages of your submission.

- Obesity is a complex condition that currently has limited treatment options.
- No products on the market that currently address appetite or satiety, so see a place for this product.
-
-
-

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

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

Thank you for agreeing to give us a statement on your view of the technology and the way it should be used in the NHS.

Healthcare professionals can provide a unique perspective on the technology within the context of current clinical practice which is not typically available from the published literature.

To help you in making your statement, we have provided a template. The questions are there as prompts to guide you. It is not essential that you answer all of them.

Please do not exceed the 8-page limit.

About you

Your name: Professor John Wilding

Name of your organisation : University of Liverpool and Aintree University Hospital NHS Foundation Trust

Are you (tick all that apply):

- a specialist in the treatment of people with the condition for which NICE is considering this technology? **Yes**

- a specialist in the clinical evidence base that is to support the technology (e.g. involved in clinical trials for the technology)? **Yes - I have extensive experience in clinical trials with various obesity pharmacotherapies, and am familiar with the technology (naltrexone / bupropion prolonged release) and the associated published clinical trial data (none of these were conducted in the UK or EU, so unlikely any UK or European investigators will have direct experience in clinical trials).**

- an employee of a healthcare professional organisation that represents clinicians treating the condition for which NICE is considering the technology? If so, what is your position in the organisation where appropriate (e.g. policy officer, trustee, member etc.)? **Yes**

- other? (please specify) **No**

Links with, or funding from the tobacco industry - please declare any direct or indirect links to, and receipt of funding from the tobacco industry: None

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

What is the expected place of the technology in current practice?

How is the condition currently treated in the NHS? Is there significant geographical variation in current practice? Are there differences of opinion between professionals as to what current practice should be? What are the current alternatives (if any) to the technology, and what are their respective advantages and disadvantages?

At present obesity and overweight are treated in primary and secondary care in the NHS, with those with less severe obesity sometimes being referred to commercially provided weight loss programmes. There is a lot of geographical variation on what is provided and who provides it. Most experts in the field (myself included) agree that there is a need to provide more comprehensive and equitable services for people with obesity. This is perhaps best summarised in the RCP document 'Obesity: comprehensive care for all' and the NICE CG189 both of which I contributed to.

Obesity treatment should include a multicomponent approach to lifestyle incorporating dietetic, physical activity and psychological support and this alone only results in 3-5kg sustained weight loss on average, with only about one-third of people achieving 5% weight loss (NICE CG189). Bariatric surgery is highly effective for selected people with more severe obesity but access is limited (less than 6000 operations per year for a population of over 1.5million eligible according to NICE guidance.

There are currently very limited pharmacological options that can bridge the gap between lifestyle approaches (as an adjunctive treatment to lifestyle) and bariatric surgery, as orlistat is the only available drug. The use of this drug is often limited by GI adverse effects and very few patients persist on therapy. There is therefore a very real and immediate need for new pharmacological options to treat obesity.

Naltrexone/Bupropion sustained release is a combination of two drugs that appear to act by both enhancing the pro-opiomelanocortin satiety signalling and also by inhibiting some reward pathways in the CNS. Clinical trials show reasonable efficacy as an adjunct to lifestyle advice in people without diabetes with about 7.8-8.1kg weight loss overall (completers) compared to 1.8 – 2.1 kg (lifestyle alone) over 6-12 months; the intention to treat analyses of these trials shows mean weight loss of 5.4 – 5.7 kg (vs 1.4-1.9kg for lifestyle). About half the patients will achieve the weight loss of 5% and about a quarter 10% weight loss, which is 2-3 x greater than lifestyle alone. Weight loss was greater in one trial where a more intensive lifestyle modification was provided. Weight loss responses seem slightly less in people with type 2 diabetes (5.9 vs 2.2kg).Overall cardiovascular risk factors blood glucose / HbA1c and lipids improved more with active treatment than placebo, however blood pressure and pulse did increase slightly (1-2mmHg, 2-3 bpm); there is some reassurance from a CV outcomes study (the LIGHT study) that does not suggest an increase in adverse cardiovascular events.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Are there any subgroups of patients with the condition who have a different prognosis from the typical patient? **Responses seem slightly less in people with diabetes, but are still clinically useful** Are there differences in the capacity of different subgroups to benefit from or to be put at risk by the technology? **No**

In what setting should/could the technology be used – for example, primary or secondary care, specialist clinics? Would there be any requirements for additional professional input (for example, community care, specialist nursing, other healthcare professionals)? **The technology could be used both in specialist clinics and in primary care where there is the capacity to provide the necessary background lifestyle support. The drug may also be useful in 'tier 3' specialist clinics where many patients are considering bariatric surgery as another option.**

If the technology is already available, is there variation in how it is being used in the NHS? Is it always used within its licensed indications? If not, under what circumstances does this occur? **It is not currently available in the NHS.**

Please tell us about any relevant **clinical guidelines** and comment on the appropriateness of the methodology used in developing the guideline and the specific evidence that underpinned the various recommendations.

The most relevant clinical guideline is NICE CG 189 which considers pharmacotherapy for obesity with orlistat. Naltrexone / Bupropion sustained release could potentially be used in the same group of people suitable for orlistat.

The advantages and disadvantages of the technology

NICE is particularly interested in your views on how the technology, when it becomes available, will compare with current alternatives used in the UK. Will the technology be easier or more difficult to use, and are there any practical implications (for example, concomitant treatments, other additional clinical requirements, patient acceptability/ease of use or the need for additional tests) surrounding its future use?

The technology seems to be straightforward to use as long as patients are properly counselled about lifestyle during therapy, warned of adverse effects and monitored with regard to response (especially with regard to stopping rules (see below), blood pressure and pulse (which should be checked before and during treatment). Patients taking opioid analgesics should not be prescribed naltrexone / bupropion (due to diminished therapeutic effect of the analgesic) and should be warned regarding over the counter use of opioid containing analgesics. The adverse effects may be less troublesome to some patients that is seen with orlistat, which is the only currently available alternative.

If appropriate, please give your view on the nature of any rules, informal or formal, for starting and stopping the use of the technology; this might include any requirements for additional testing to identify appropriate subgroups for treatment or to assess response and the potential for discontinuation.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

The most important stopping rule is the one that is in the spc that requires stopping the medicine if a 5% weight loss has not been achieved after 12 weeks at the target dose (this is effectively after 16 weeks of therapy due to the need for dose titration. This is important as it means that non-responders to the drug should not be exposed for long periods of time unnecessarily and evidence provided in analysis of the trials suggests that responders will achieve a greater mean weight loss (ie greater clinical benefit) than those who are non responders (mean weight loss in responders is over 11kg).

If you are familiar with the evidence base for the technology, please comment on whether the use of the technology under clinical trial conditions reflects that observed in clinical practice. Do the circumstances in which the trials were conducted reflect current UK practice, and if not, how could the results be extrapolated to a UK setting? What, in your view, are the most important outcomes, and were they measured in the trials? If surrogate measures of outcome were used, do they adequately predict long-term outcomes?

Athough the clinical trials were conducted in North America they included a range of people from different ethnic backgrounds and with a range of co-morbidity and from experience the clinical responses are likely to be similar in the UK. Clinical trials often have inclusion and exclusion criteria, that restrict the populations studied, but from the published evidence the people included would be reasonably representative of the UK population requiring treatment for obesity. The main outcomes other than body weight are changes in CV risk factors and diabetes control that were measured in the trials. It is important to note that there was also an evaluation of quality of life included in the trials and overall this improved to a greater extent than was seen with lifestyle alone.

What is the relative significance of any side effects or adverse reactions? In what ways do these affect the management of the condition and the patient's quality of life? Are there any adverse effects that were not apparent in clinical trials but have come to light subsequently during routine clinical practice?

The main adverse events are related to the known pharmacology of the two drugs and include nausea, dry mouth, constipation, dizziness, headache and sometimes vomiting. These are generally manageable although it should be noted that some patients (about 23.9% in the clinical trials) withdrew from treatment due to adverse effects (compared to 11.9% with placebo). It should also be noted that as naltrexone is an opioid receptor antagonist the drug should not be used in combination with opioid analgesics (it may reduce the therapeutic effect). From a cardiovascular perspective (which has been a problem with some previous obesity drugs) there do not seem to be major concerns and the CV outcomes study (the LIGHT study) provides some reassurance in that regard.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

Equality and Diversity

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. Please let us know if you think that this appraisal:

- Could exclude from full consideration any people protected by the equality legislation who fall within the patient population for which [the treatment(s)] is/are/will be licensed; **No**
- Could lead to recommendations that have a different impact on people protected by the equality legislation than on the wider population, e.g. by making it more difficult in practice for a specific group to access the technology;**No**
- Could lead to recommendations that have any adverse impact on people with a particular disability or disabilities; **No**

Please tell us what evidence should be obtained to enable the Committee to identify and consider such impacts

Any additional sources of evidence

Can you provide information about any relevant evidence that might not be found by a technology-focused systematic review of the available trial evidence? This could be information on recent and informal unpublished evidence, or information from registries and other nationally coordinated clinical audits. Any such information must

include sufficient detail to allow a judgement to be made as to the quality of the evidence and to allow potential sources of bias to be determined.

I am not aware of any evidence other than the published clinical trials which I am sure the committee will already be aware of.

Implementation issues

The NHS is required by the Department of Health and the Welsh Assembly Government to provide funding and resources for medicines and treatments that have been recommended by NICE technology appraisal guidance. This provision has to be made within 3 months from the date of publication of the guidance.

If the technology is unlikely to be available in sufficient quantity, or the staff and facilities to fulfil the general nature of the guidance cannot be put in place within 3 months, NICE may advise the Department of Health and the Welsh Assembly Government to vary this direction.

Please note that NICE cannot suggest such a variation on the basis of budgetary constraints alone.

How would possible NICE guidance on this technology affect the delivery of care for patients with this condition? Would NHS staff need extra education and training? Would any additional resources be required (for example, facilities or equipment)?

I do not think any additional resources or training would be required other than a modest amount of learning for prescribers as would be appropriate for any new medication.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single Technology Appraisal (STA)

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Naltrexone-bupropion for managing overweight and obesity

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 09/03/2017

Source of funding: This report was commissioned by the NIHR HTA Programme as project number STA 14/55/01.

Declared competing interests of the authors
None.

Acknowledgements
None.

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:
Fayter D, Ramaekers BLT, Grimm S, Armstrong N, Noake C, Wei C-Y, Worthy G, Carrera S, Riemsma R, Joore MA, Kleijnen J. Naltrexone-bupropion for managing overweight and obesity: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2017.

Contributions of authors
Rob Riemsma acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Bram Ramaekers acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Sabine Grimm, Nigel Armstrong and Ching-Yun Wei acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Debra Fayter and Sofia Carrera acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Caro Noake critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

ACM	All-cause mortality
ACMM	Adjusted censored mixture model
AE	Adverse Events
BCOF	Baseline observation carried forward
BMI	Body mass index
BNF	British National Formulary
BP	Blood pressure
BI	budget impact
BMOD	Behaviour modification
BSC	Best supportive care
CE	Cost effectiveness
CEA	Cost effectiveness Analysis
CEAC	Cost effectiveness acceptability curve
CHD	Coronary heart disease
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence Interval
COE	Control of Eating questionnaire
COR	Contrave obesity research
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company's submission
CSR	Clinical study report
CVD	Cardiovascular disease
CVOT	Cardiovascular outcomes trial
DM	Diabetes mellitus
EMA	European Medicines Agency
EPAR	European public assessment report
EQ-5D	European Quality of Life-5 Dimensions
ERG	Evidence Review Group
EUR	Erasmus University Rotterdam
FCI	Food Craving Inventory
FDA	Food and Drug Administration
GP	General practitioner
HOMA-IR	Homeostasis model assessment of insulin resistance
HR	Hazard ratio
HRQL	Health Related Quality of Life
hs-CRP	High-sensitivity C reactive protein
HTA	Health Technology Assessment
IC	Indirect Comparison
ICER	Incremental Cost Effectiveness Ratio
IDS-SR	Inventory of Depressive Symptoms – Subject related
ITC	Indirect treatment comparison
ITT	Intention to Treat
IWQOL-Lite	Impact of Weight on Quality of Life-Lite version
KSR	Kleijnen Systematic Reviews
LOCF	Last observation carried forward
LYS	Life Year Saved
MACE	Major adverse cardiovascular events
MD	Mean Difference
MeSH	Medical Subject Headings
MHRA	Medicines and Healthcare Products Regulatory Agency
mITT	Modified ITT analysis
MTC	Mixed Treatment Comparison

NA	Not applicable
NB32	Naltrexone 32mg plus bupropion 360mg
NHS	National Health Services
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not Reported
OR	Odds Ratio
ORL	Orlistat
OS	Overall survival
PRESS	Peer Review of Electronic Search Strategies
PSA	Probabilistic Sensitivity Analyses
PSS	Personal Social Services
QALY(s)	Quality-adjusted Life Year(s)
QoL	Quality of life
RCT	Randomised Controlled Trial
RR	Relative Risk; Risk Ratio
SAE	Serious Adverse Events
SD	Standard deviation
SM	Standard management
SPC	Summary of product characteristics
STA	Single Technology Appraisal
UMC	University Medical Centre
T2DM	Type 2 Diabetes Mellitus
TEAEs	Treatment-emergent adverse events
TESAEs	Treatment-emergent serious adverse events
TIA	Transient Ischaemic Attack
UK	United Kingdom
VAS	Visual Analogue Scale
WHO	World Health Organisation

Table of Contents

Table of Tables

Table of Figures

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

The NICE scope describes the decision problem as naltrexone-bupropion prolonged release (32mg daily) or NB32 for managing overweight (≥ 27 kg/m2 to < 30 kg/m^2; in the presence of one or more weight-related co-morbidities) and obesity (≥ 30 kg/m^2). The comparators are described as: standard management without naltrexone-bupropion and orlistat (360 mg/day). Standard management is not defined in the NICE scope.

NB32 is indicated, as an adjunct to a reduced-calorie diet and increased physical activity, for the management of weight in adult patients (≥18 years). NB32 treatment should be discontinued after 16 weeks if patients have not lost at least 5% of their initial body weight. Likewise, treatment with orlistat should be stopped after 12 weeks if patients have been unable to lose at least 5% of their body weight since the start of treatment. In most trials NB32 and orlistat are continued throughout the trial, usually one year's duration.

The main question regarding the decision problem is the appropriateness of the intervention and comparator insofar as what constitutes standard management in clinical practice. The company assumed that standard management was more like that in the COR-I and COR-II trials in which patients received advice on diet and exercise but were not allowed to participate in a weight loss programme. However, if those who are prescribed NB32 or orlistat would engage in a concomitant weight loss programme then the intervention might be more like that, referred to as 'intensive behaviour modification', in the COR-BMOD trial. Similarly, if those who are eligible for either NB32 or orlistat would otherwise engage in a weight loss programme then the comparator might be more like that in COR-BMOD.

1.2 Summary of clinical effectiveness evidence submitted by the company

The company's submission included data from four main trials comparing NB32 to placebo as an adjunct to standard management: COR-I, COR-II (general overweight and obese population), COR-BMOD (intensive behaviour modification) and COR-DM (diabetes population). Mean BMI across the trials was 36 to 37. Approximately 2% of participants were overweight and 98% obese.

All trials were multicentre and all were conducted in the US. All had a joint primary outcome of percentage change in total body weight and proportion of patients with >5% decrease in total body weight. Three trials measured outcomes at week 56. One trial, COR-II measured the primary outcome at 28 weeks. In COR-II, NB32 patients who had lost less than 5% of their body weight at visits between weeks 28 and 44 were re-randomised to continue with NB32 or escalate to NB48. The four trials included 4,536 patients. Of these 2,510 patients were randomised to NB32, 578 to NB16 (in COR-I) and 1,448 randomised to placebo.

The main results presented in the company submission (CS) were based on a modified intention-to-treat (mITT) analysis. According to the CS this was defined as "*all randomised patients with a post-baseline body weight measurement obtained while the patient remained on study medication."* In this modified ITT analysis, approximately 20% of randomised patients were not included in the analyses.

Direct evidence from the four main trials showed that the mean difference in percentage weight change at week 56 from baseline was -3.3 (95% CI: -4.3 to -2.2) for COR-DM, favouring NB32 compared with placebo; -4.2 (95% CI: -5.6 to -2.9) for COR-BMOD; -4.6 (95% CI: -5.2 to -3.9) for COR-II (at 28 weeks); and -4.8 (95% CI: -5.6 to -4.0) for COR-I. Analyses for the number of patients with ≥ 5% decrease in weight at week 56 also significantly favoured NB32 over placebo in all four trials (Odds

ratios: 3.4 (95% CI: 2.2 to 5.5) for COR-DM; 2.9 (95% CI: 2.0 to 4.1) for COR-BMOD; 6.6 (95% CI: 5.0 to 8.8) for COR-II (at 28 weeks); and 4.9 (95% CI: 3.6 to 6.6).

The percentages of overweight patients (BMI < 30 kg/m^2) in the trials are too small to present meaningful subgroup analyses.

Adverse events occurred in 83.1% to 93.7% of treatment groups and 68.5% to 88.0% of placebo groups. Approximately 58% to 76% of these were attributed to the drug in NB32 groups across the trials. Serious adverse events occurred at similar rates in treatment and placebo groups across the trials. However, a larger number of patients discontinued due to adverse events across the trials (19.5% to 29.4% for treatment groups) versus 9.8% to 15.4% in placebo groups).

The main category of adverse event occurring more frequently in treatment groups across the trials was gastrointestinal disorders. Nausea, in particular, occurred frequently and more often in treatment groups. Across the trials, rates of nausea ranged from 29.2% to 42.3% in treatment groups. Rates ranged from 5.3% to 10.5% in placebo groups. Vomiting, constipation and dry mouth also occurred more frequently in treatment groups although at a lower rate than that of nausea. Nervous system disorders such as headache, dizziness and tremor occurred more frequently in treatment groups.

The incidence of events of particular concern (serious cardiovascular disorders and suicidality measured on IDS) was extremely small and any differences between groups could not be ascertained in view of the small numbers in both groups.

No trials were identified that compared NB32 directly with orlistat or with different types of behavioural interventions. Therefore, the company performed indirect comparisons to compare NB32 with orlistat using placebo as the common comparator. Twenty trials were included in the indirect treatment comparison (ITC), four for NB32 and 16 for orlistat.

Results for mean percentage weight change from baseline at one year showed that there were no significant differences between NB32 and orlistat for people with diabetes and for all patients combined. There was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients are excluded (MD 1.13 (95% CrI: 0.44, 1.80)). The difference is most significant for the third sensitivity analysis, where studies with 'intensive' behaviour modification (BMOD and XENDOS) were also excluded (MD 2.98 (95% CrI: 1.60, 4.36)).

Results for ≥5% reduction in weight at one year showed that there were no significant differences between NB32 and orlistat for people with diabetes and for all patients combined. There was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients are excluded (OR 0.77 (95% CrI: 0.61, 0.96)). The difference is most significant for the third sensitivity analysis, where studies with 'intensive' behaviour modification (BMOD and XENDOS) were also excluded (OR 0.44 (95% CrI: 0.23, 0.84)).

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

The company submission (CS) and response to clarification provided sufficient details for the ERG to appraise the searches for eligible trials. Searches were carried out in accordance with the NICE guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4 using a good range of databases. Additional searches of conference proceedings and organisational websites were reported, along with the checking of reference lists of existing systematic literature reviews (SLRs) and meta-analyses.

The four main trials comparing NB32 to placebo are of high quality. However, there are a number of limitations when applying them to clinical practice. There are very little data on ethnic groups relevant

to the UK (particularly people from Asia) within the NB32 trials, therefore it is not possible to make any firm conclusions for that group. There are very few overweight as opposed to obese participants in the trials. The majority of the participants in the NB32 trials are female. Trials do not measure weight loss beyond 56 weeks. The large dropout from the NB32 trials (up to 50%) is relevant to practice. The US setting may reflect a different patient profile and differing approaches to standard care than in a UK setting.

A comparison between NB32 (plus standard management) versus intensive behaviour modification is missing. Furthermore, comparisons between NB32 and orlistat are based on indirect comparisons only.

The company used modified ITT data from NB32 trials, but this is misleading. The mITT population in the NB32 trials is very different from mITT populations in the orlistat trials. In the NB32 trials, 21.9% of patients receiving NB32 were randomised but excluded from the analyses against 1.6% of patients receiving orlistat.

Comparison with orlistat may be biased in favour of NB32. NB32 trials were published in 2010 or later; most of the trials with orlistat were published before 2005, so caution should be exercised when making indirect comparisons; this is particularly true for conditions such as diabetes where background standard therapy (for glucose and lipids especially) may be very different now.

We have reproduced the company's indirect analyses comparing orlistat and NB32 using full ITT data from the NB32 trials. The results show that the positive effects of NB32 when compared to orlistat have all disappeared. For the first outcome (≥5% reduction in weight at one year), there was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded using mITT data. However, in both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: OR = 1.09 (95% CrI: 0.87 to 1.36), ITT-BOCF: OR = 1.06 (95% CrI: 0.84 to 1.33). Moreover, although none of the differences are statistically significant, all results now favour orlistat.

For the second outcome (mean percentage weight change at one year), using mITT data there was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded. However, in both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: MD= -0.09 (95% CrI: -0.77 to 0.58), ITT-BOCF: MD = -0.54 (95% CrI: -1.21 to 0.12). Moreover, although most of the differences are not statistically significant, most results now favour orlistat.

Standard management in the UK might be better reflected by COR-BMOD; therefore, we have included a new analysis: an indirect comparison of NB32 plus intensive behaviour modification (COR-BMOD) versus orlistat plus intensive behaviour modification (XENDOS). The results show that both outcomes significantly favour orlistat over NB32 (≥5% reduction in weight at one year: OR 1.86 (95% CI: 1.30 to 2.66); Mean percentage weight change from baseline (CFB) at one year: MD -2.09 (95% CI: -3.53 to -0.65)).

Finally, we performed our preferred analyses, i.e. using full ITT data and no pooling of NB32 trials (using only COR-I ITT data for non-diabetics, instead of COR-I, COR-II and COR-BMOD combined). The results for 'obese patients with T2DM' and 'intensive behaviour modification' are the same as before, but results for 'obese patients without T2DM' have changed considerably again, and are almost the same as in the company's original analyses. Both outcomes show no significant difference between NB32 and orlistat, but both favour NB32.

The table below shows the main results for obese people with diabetes, obese people without diabetes and NB32 plus intensive behaviour modification versus orlistat plus intensive behaviour modification.

Table 1.1: Company results versus ERG results

Population		Company analyses (mITT data)*	Company analyses (ITT-BCFA data)**	ERG preferred analyses**
		Orlistat vs NB32	Orlistat vs NB32	Orlistat vs NB32
Obese people with T2DM				
≥5% reduction in weight at 1 year	OR	1.09 (0.63 to 1.88)¶	1.59 (0.89 to 2.79)¶	1.59 (0.89 to 2.79)¶
Mean % weight CFB at 1 year	MD	0.21 (-0.87 to 1.30)†	-1.21 (-2.30 to -0.11)¶	-1.21 (-2.30 to -0.11)¶
Obese people without T2DM				
≥5% reduction in weight at 1 year	OR	0.77 (0.61 to 0.96)†	1.06 (0.84 to 1.33)¶	0.61 (0.31 to 1.22)†
Mean % weight CFB at 1 year	MD	1.13 (0.44 to 1.80)†	-0.54 (-1.21 to 0.12)¶	1.11 (-0.39 to 2.63)†
Intensive behaviour modification				
≥5% reduction in weight at 1 year	OR	1.22 (0.84 to 1.77)¶	1.86 (1.30 to 2.66)¶	1.86 (1.30 to 2.66)¶
Mean % weight CFB at 1 year	MD	-0.21 (-1.28 to 1.70)¶	-2.09 (-3.53to -0.65)¶	-2.09 (-3.53to -0.65)¶

Results are OR with 95% CI/CrI for ≥5% reduction in weight at 1 year and mean difference (MD) with 95% CI/CrI for mean % weight CFB at 1 year.
An OR less than one favours NB32 over orlistat and a CI including 1 is not significant. A MD of >0 favours NB32 over orlistat and indicates greater % weight reduction and a CI including 0 is not significant.
¶ = Favours orlistat; † = Favours NB32.
*) Bayesian NMA (OR, 95% CrI) using mITT data; **) Using the Bucher method for indirect comparisons and ITT-BCFA data.
FE = fixed effect; ITT-BCFA = all randomised patients with baseline-carried-forward analysis; MD = Mean Difference; mITT = modified intention-to-treat analysis; NB32 = naltrexone 32mg plus bupropion; OR = Odds Ratio; T2DM = Type 2 diabetes mellitus;

Which of the estimates of treatment effect is more applicable to clinical practice depends on the definition of standard management. If individuals who are eligible for NB32 would also engage in a weight loss programme when prescribed NB32 then the so-called intensive behaviour modification estimate might be more applicable. If this is not the case, then an estimate excluding intensive behaviour modification might be more appropriate. Of course, the estimate of 1.06 (0.84 to 1.33) is based on pooling both the trials with and without intensive behaviour modification and it is therefore tempting to infer that this represents clinical practice, where some do and some do not engage in weight loss programmes. This must be regarded with caution for a number of reasons, which include uncertainty as to the precise proportion who would engage in a weight loss programme and the degree of resemblance between such a programme and the intensive behaviour modification in COR-BMOD.

1.4 Summary of cost effectiveness evidence submitted by the company

The company conducted systematic reviews to identify relevant cost effectiveness studies, health-related quality of life (HRQoL) studies, resources and costs studies. The company did not identify any study investigating the cost effectiveness of NB32 adjunct to standard non-pharmacological

management in the population of interest for the current decision problem, and hence developed a *de novo* model with a lifetime horizon.

The company developed an economic model using an individual-level approach, more specifically a discrete event simulation (DES). It was argued that an individual-level approach is better suited than a cohort-level approach to capture the chronic implications of both weight and weight-related health events in a heterogeneous group of overweight and obese patients. The DES model was implemented in Excel using the "discretely integrated condition event" (DICE) principles and structure. The company used an economic evaluation by Ara et al. (also an individual-level model) as a starting point, which is from a 2012 Health Technology Appraisal comparing different pharmacological treatments for obesity. The following events are considered in the economic model:

- treatment discontinuation;
- development of T2DM;
- first cardiovascular event (either stroke or MI);
- second cardiovascular event (either stroke or MI) and;
- death.

Upon model entry, a simulated patient is assigned a profile of sampled baseline characteristics that are explanatory factors for risks, costs and/or utility in the model (sampled baseline characteristics as well as random numbers for the sampled patient are equal across all three treatments). The baseline profile characterises the individual patients by:

- age (years);
- gender (male, female);
- height (meters);
- BMI (kg/m^2);
- T2DM status (yes, no);
- smoker status (current, previous, never);
- receive insulin, if diabetic (yes, no);
- receive statins (yes, no);

The company stated that the economic analysis aimed to reflect the patient group for which the drug is licensed: adult patients who are obese (BMI ≥30kg/m2), or overweight (BMI ≥27kg/m2 and <30kg/m2) in the presence of one or more weight-related comorbidities (e.g., T2DM, dyslipidaemia, or controlled hypertension). The company assumed that no patients would have a history of angina or diabetes other than T2DM and no patients received anti-hypertensive medication and/or aspirin.

In line with the final scope and licensed indications, the company considered orlistat as an adjunct to standard management and standard management alone as comparators for NB32 as an adjunct to standard management. NB32 is implemented as per its European Medicines Agency (EMA) Summary of Product Characteristics (SmPC) posology and method of administration, incorporating a four week escalation period, after which the maximum recommended daily dose of 32mg naltrexone hydrochloride and 360mg bupropion hydrochloride is assumed. Orlistat is similarly implemented as per its EMA SmPC posology and method of administration, a 360mg daily dose. The company specified standard management as implemented in the analysis to reflect the non-pharmaceutical dietary and lifestyle management treatment received in UK NHS practice.

Treatment effectiveness estimates (i.e. time to treatment discontinuation data, proportion of responders, and change in body weight) were mainly derived from the COR trial programme, including the COR-

I, COR-II, COR-BMOD and COR-DM trials. All the analyses were based on the company's modified ITT analysis, which reflects only those patients who have a post-baseline measurement whilst on the study drug. Time to treatment discontinuation was estimated based on the COR trial programme and extrapolated after one year using the NB-CVOT study. All patients were assumed to have discontinued after treatment duration data were unavailable in this study. It should be noted that the company used the same time to treatment discontinuation Kaplan-Meier curves for both NB32 and orlistat. The company justified this by stating that data were lacking for orlistat. Both the proportions of responders and change in body weight were obtained from the COR trial programme for NB32 and standard management; an ITC was used to calculate this for orlistat. The changes in body weight were used to predict development of T2DM, cardiovascular event (either stroke or MI) and death using parametric time-to-event models (Weibull distribution) retrieved from the report by Ara et al. Also the natural history of BMI model, to predict BMI over time, was retrieved from this report. The company stated that it was unable to make trial data comparisons of AEs associated with NB32 and orlistat because details from clinical literature and regulatory documents on orlistat were insufficient. Therefore, the company assumed equal AE related costs for NB32 and orlistat. The impact of AE on utility scores was not incorporated by the company.

The company applied a Tobit model to estimate utility values based on the Public Health England weight management economic assessment tool v2 (Health Survey for England EQ-5D data analysis). This model includes explanatory variables for BMI, age, gender, and obesity-related conditions (stroke, MI, cancer and T2DM).

Costs in the model consisted of drug acquisition costs, non-drug costs related to standard management (applicable to all treatments considered), obesity-related comorbidity costs and adverse event costs. Drug acquisition costs for NB32 and orlistat were based on the list price and Monthly Index of Medical Specialities respectively. The non-drug resource use items comprising standard management in the model consisted of GP visits, nurse visits and blood tests which were informed by the COR trials, literature and clinical opinion. Moreover, obesity related comorbidity costs were retrieved from the literature and for AE the costs of one GP visit were assumed.

The company's model uses 1,000 patient profiles for their deterministic analysis. For the probabilistic sensitivity analysis (PSA), the company used only 500 patient profiles and 100 PSA simulations. Moreover, not all model parameters were incorporated in the PSA.

In the base-case deterministic analysis, NB32 was associated with an incremental QALY gain of 0.0765 QALYs versus standard management, and 0.0192 QALYs versus orlistat. The incremental costs of NB32 were £1,044 versus standard management and £750 versus orlistat. The incremental cost effectiveness ratio (ICER) of NB32 versus standard management was £13,647 per QALY. The estimated ICER versus orlistat was £32,084 per QALY. Subgroup analyses performed by the company indicated that the ICERs of NB32 compared with standard management and orlistat were £5,059 and £72,069 respectively for T2DM patients and £6,283 and £28,291 respectively for non-T2DM patients.

The deterministic sensitivity analyses performed by the company showed that the most influential parameters were the parameters of the Tobit model for utilities and the discount rate for QALYs, as well as parameters related to the measures of relative efficacy from the ITC. The company performed scenario analyses on the following model aspects: the time period over which weight is regained, the cost of T2DM, the utility estimates, costs of AEs, discounting, and the time horizon. The most influential scenarios were shortening the time period for weight regain from three to two years (ICER £41,016), and shortening the time horizon from lifetime to 15 years (£53,514).

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The majority of the cost effectiveness searches in the CS were well documented and easily reproducible, and were carried out in line with the NICE guide to the methods of technology appraisal.

The ERG considered it reasonable to use the economic model by Ara et al. as a starting point for the current analysis. It should be noted that the company deviated from the assumption made by Ara et al., that patients would have regained weight to obtain the *baseline* BMI within three years in a linear fashion and assumed instead that patients would have regained weight to obtain the *age/sex predicted* BMI in three years. The company did not provide justification for why their deviation from Ara et al.'s assumption was 'logical' and plausible. Hence, to be consistent with Ara et al., the ERG preferred to assume weight regain to the baseline BMI in its base-case. Furthermore, the linear weight regain over the time-span of three years was implemented incorrectly in the model in that the weight regain occurs instantaneously at the end of the three year period. The ERG also questioned the (justification for the) assumption of equivalent weight loss at similar assessment times. The company's model assumed weight loss for orlistat patients at weeks 12 and 52 to be comparable to weight loss for NB32 patients at weeks 16 and 56. This was not justified besides stating that this assumption was also upheld within the ITC. The model only includes the possibility of two subsequent cardiovascular events (i.e. either two strokes, two MIs or one stroke and one MI), implicitly assuming that the impact of the third cardiovascular event on costs, quality of life and survival, is negligible. It can, however, be questioned whether having a stroke after having experienced two MIs is indeed unimportant.

The population aimed to reflect the scope. However, patient characteristics in the model were sampled from estimates that were based on a variety of sources. It is questionable whether this is reflective of UK clinical practice. The ERG agrees with using the COR trial programme patient-level data to inform baseline patient characteristics in the model (as done for age, gender and height). This follows from a) that the effectiveness estimates are derived from this population and b) that the company stated, based on clinical opinion, that patient characteristics in the COR trial programme are a fair reflection of the typical patient group that would receive NB32 in UK NHS clinical practice. However, the appropriateness of other baseline characteristics is less clear. The ERG considered the BMI sampled in the model and compared it with the baseline BMI in the COR trial programme and concluded that baseline BMI is vastly underestimated in the economic model. This is also reflected in the average baseline weight of 92kg in the model, while the averages ranged between 99kg and 105kg in the COR trial programme. Given that BMI is included as a predictive factor for utility, T2DM, cardiovascular events and death, the utility values and the time to these events in the model are overestimated, likely inducing bias in favour of NB32.

Other baseline characteristics are also potentially underestimated:

- Proportion of current smokers
- Proportion of patients receiving anti-hypertensive medication
- Proportion of patients with a history of angina and/or diabetes other than T2DM
- Proportion of patients receiving aspirin

In contrast to the above, the proportion of patients receiving statins and patients with T2DM might have been overestimated. Moreover, correlations between covariates were not incorporated in the sampling of the patient characteristics, leading to counter-intuitive patient profiles. For instance, based on the patient characteristics of the COR-I, COR-II, COR-BMOD and COR-DM trials, it becomes clear that the patients without T2DM (COR-I, COR-II and COR-BMOD trials) have different patient characteristics (e.g. regarding age, sex, hypertension status and statin use) than patients with T2DM (COR-DM trial). This is neglected in the sampling of the patient population. To address these issues,

the ERG adjusted the baseline characteristics used in the model. This included calibrating the natural history model to predict BMI over time.

The company did assume no patients had a history of angina and/or diabetes other than T2DM. This assumption was made as no data were identified on these characteristics for overweight/obese patients. The ERG agrees with this statement and would therefore argue that it can be questioned whether the results of the economic analyses are representative for patients with a history of angina and/or diabetes other than T2DM.

One major limitation of the model is the inability to incorporate re-treatment, behaviour modification treatment and/or bariatric surgery (for which patients become eligible over time once their BMI is/increases to >40kg/m2 in the model).

The ERG considers that the use of the ITT population (instead of the mITT) to inform treatment response and weight loss would have been both more appropriate and more conservative. Using the true ITT data, NB32 would achieve a smaller mean percentage of weight loss and smaller proportion of responders compared to the mITT data. It is also the ERG's view that it was inappropriate to pool from all COR studies, including COR-BMOD and COR-II. Effectiveness estimates derived from the COR-BMOD trial where NB32 was administered in combination with intensive behavioural modification are substantially different when compared to effectiveness estimates derived from studies in which NB32 was administered together with standard management only. Likewise, the ERG considers the use of COR-II for the derivation of treatment effectiveness beyond 28 weeks as inappropriate because NB32 participants with <5% weight loss at visits between Weeks 28 and 44 were re-randomised. The ERG therefore considers that NB32 treatment effectiveness estimates should only be derived from the COR-I and COR-DM trials.

Because of the following reasons, the ERG believes time to discontinuation (TTD) is underestimated for all treatments in the model but in particular for orlistat:

(1) TTD estimates for the period after the one year assessment were derived from the NB-CVOT study in which patients had characteristics associated with an increased risk of CV outcomes, potentially leading to a shorter TTD.
(2) The end of the NB-CVOT study was used as the maximum TTD, whether patients in that study had discontinued or not.
(3) The company claims that the most reasonable and conservative assumption was to assume that TTD for orlistat would follow a similar trajectory to NB32, given that patient-level data for orlistat were unavailable. However, the ERG found publications reporting TTD for orlistat, which reveal that orlistat TTD was longer than the 12.29 months estimated by the model, with many studies reporting that the proportion of patients still receiving orlistat at 12 months was >50%.
(4) For the derivation of the orlistat TTD, the KM estimates for NB32 TTD for the first 16 weeks were linearly scaled to fit the first 12 weeks of orlistat treatment.

The ERG considers the company's claim that not accounting for a HRQoL impact of AEs in the economic model is conservative as highly questionable. The company provided no systematic overview of evidence that showed that the AE profile of orlistat was indeed worse than that of NB32. There is no direct evidence comparing the two drugs and indirect treatment comparisons between the drugs focused on efficacy but not on safety outcomes. Therefore the company's assertion of the likely superiority of NB32 in relation to orlistat in terms of AE remains speculative. Upon request, the company provided a scenario analysis in their response to clarification question B13, in which "pragmatic application of on-

treatment disutilities has been provided", assuming all AEs to be associated with a utility decrement of 0.05 for the duration of one week. This analysis increased the company's base-case ICERs against orlistat and SM by £188 and £87 per QALY gained, respectively.

The ERG is concerned that the regression model that informs the utility estimates does not appear to be published in a peer-reviewed journal. As a consequence, given the limited amount of details, the validity of these regression models to estimate utility values cannot be assessed by the ERG. However, upon request from the ERG, the company assessed the face validity of the utility estimates. The company stated that the utility values predicted by the Tobit model for the healthy population resembled the ones from the general UK population and that the remainder of the predicted utilities lay below these, demonstrating face validity.

The ERG considered it plausible to use Ara et al. to inform healthcare resource use assumptions. Regarding the costs of standard management, it is unclear to the ERG why the company added a GP visit for the 52 week assessment for patients receiving standard management only. Therefore, the ERG removed this GP visit for patients receiving standard management only.

The ERG ran the deterministic CS base-case model with 1,000 individual sampled patients, which resulted in an ICER of NB32 versus orlistat ~£3,000 higher than the base-case results reported in the CS. In the ERG's further analyses, there was substantial variation in the ICERs obtained in model runs when a different set of random numbers was used and a new set of patients were sampled. Based on the ERG's findings, and the uncertainty that the company's diagnostic exercises truly reflected the stability of the model, the ERG believes that the model should ideally be evaluated using a much larger number of sampled patients (more than the 1,000 that are used in the CS base-case). However, model run times were prohibitive (six hours on average per model run with 1,000 patient profiles) and the model was restricted to incorporate a maximum of 1,000 patients. Moreover, the ERG believes the PSA results in the CS are flawed for multiple reasons: 1) the low number of individual sampled patients (500) included in the PSA; 2) the low number (100) of PSA simulations and; 3) the exclusion of key input parameters from the PSA (e.g. TTD, natural history of BMI model, obesity-related events).

The structure and technical implementation of the company's model caused long run times (6 hours on average), and caused the model to crash on multiple computers. This hampered the company's and the ERG's ability to perform an appropriate PSA and the ERG's ability to check the model's validity and perform further scenario analyses (other than those that were described below). It should be considered whether simpler approaches (e.g. an individual-level state transition model) would have been more appropriate to reflect this decision problem, given the gain in transparency and that it would have been possible to reflect the condition-specific events in such a model. An individual-level state transition approach would potentially resolve most of the validity issues (e.g. the fact that BMI was not accurately reflected at each time period).

The ERG considered the internal validity of the model (e.g. checking formulae in the DICE sheet, examining the implementation of TTD in the model, examining available intermediate outcomes). However, the ERG was unable to examine the internal validity of the model according to its usual standards. This was mainly a consequence of the long model run times for one single deterministic analysis (six hours) and the inability to examine intermediate outcomes. For instance, the nature of the model hampered the ERG's ability to do sensitivity analysis; extreme value analysis; trace analysis/analysis of intermediate outcomes which are recommended by the ISPOR taskforce on model transparency and validation. Therefore, the ERG wishes to note that it cannot be guaranteed that there are no modelling errors (in addition to the methodological flaws described below). In this light, the

ERG considers it troublesome that the company did not provide the results of the internal validation it performed (as requested in response to clarification question B19).

One of the main validity issues or methodological flaws the ERG encountered was the inaccurate reflection of patients' BMI and consequently health-related quality of life. After the first year, patients have on average only three events in 32.8 years, equalling to an average of one event per 10.6 years. This entails that BMI after the first year is only updated on average once every 10.6 years (implicitly assuming a stable BMI in the periods between events), while this should be updated at least annually to reflect the increasing BMI due to its correlation with age (as reflected in the natural history model predicting BMI over time). This could have been solved by an annual updating event, the integration of the BMI function or the use of a different model structure. Apart from the annual updating or integration of BMI (and the impact on associated risks and utility values), the lack of model updating also affects other assumptions in the model. For example, the assumption regarding weight regain after treatment discontinuation for NB32 and orlistat was intended to reflect linear weight regain for a period of three years after which the BMI is obtained (predicted by the natural history model). However, if there is no event in this three year weight regain period, which is more likely than not (based on the average of one event per 10.6 years), the BMI estimated at the time of treatment discontinuation is maintained for this weight regain period of three years after which the weight is regained instantly. It should be noted that, if the death event were to be excluded from this calculation, the average time until one event would increase to 17.2 years. According to the ISPOR taskforce on DES, it would have been recommended to incorporate 'time checks' (i.e. 'update events'). Given that BMI is underestimated as a consequence of this methodological flaw, the utility values and the time to the events in the model are overestimated, likely inducing bias in favour of NB32. Moreover, assuming stable BMI for long periods of time also limits the face validity of the model.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

The majority of searches for eligible studies in the CS were well documented and easily reproducible. Searches were carried out on a good range of databases. The strategies utilised recognised study design filters. Supplementary searches of conference proceedings and organisational websites, and the checking of references lists were undertaken by the company in order to identify additional studies not retrieved by the main searches.

Four good quality large RCTs for NB32 and 16 comparator trials were included in the submission. Analyses were presented for all patients and people with and without T2DM, including a large number of sensitivity analyses.

The economic model structure is similar to the assessment by Ara et al., which is a Health Technology Appraisal report (2012) comparing different pharmacological treatments for obesity.

1.6.2 Weaknesses and areas of uncertainty

There were limitations with the use of indexing terms on Embase.com searches, as strategies only used EMTREE. Although some mapping between indexing terms does take place on Embase.com it is possible that relevant MEDLINE indexing terms (MeSH) will not have been included in the search, and potentially relevant records could have been missed

The main weakness of the CS is the use of mITT populations for the NB32 trials. These data overestimate the benefits of NB32 over placebo or orlistat when compared to the true ITT data.

Uncertainty remains surrounding the effectiveness of NB32 for patients who are overweight with comorbidities as opposed to obese; ethnic groups relevant to a UK setting and those who have previously used orlistat. Further uncertainties include any further weight loss and maintenance of weight loss after 56 weeks, and retreatment with NB32. The relative benefit of NB32 in comparison to orlistat is uncertain when all data are taken into account. The benefit of NB32 when compared to an optimally delivered intensive intervention in practice is unclear as is NB32 treatment discontinuation in clinical practice.

The interpretation and validity of the results are particularly hampered given that the company's model did underestimate TTD, did not incorporate behaviour modification interventions (e.g. weight loss programmes), bariatric surgery and re-treatment nor an updating event or integration of the BMI function that was required to accurately reflect patients' expected quality of life and costs associated with resource use. The lack of an updating event or integration of BMI could significantly bias the results in favour of NB32. The model structure and technical implementation of the model hampered the assessment of validity of all parts of the model in the given time-frame. It should be considered whether simpler approaches (e.g. an individual-level state transition model) would have been more appropriate to reflect this decision problem, given the gain in transparency and given that it would have been possible to reflect the condition-specific events in such a model. An individual-level state transition approach would potentially resolve most of the validity issues (e.g. the lacking updating event).

Furthermore, the ERG considers the model as unfit for purpose, due to its extremely long run times, the fact that it crashes on many computers, and the inability to perform PSA.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

Numerous issues were identified by the ERG. The ERG was able to adjust/correct some of these issues in its base-case. The ERG base-case ICERs (deterministic) of NB32 compared with standard management and orlistat ranged between £9,813-£10,510 and £38,871-£45,694 per QALY gained respectively. Subgroup analyses performed conditional on the ERG base-case, indicated that the ICERs (deterministic) of NB32 compared with standard management and orlistat were £10,535 per QALY gained and dominated respectively for T2DM patients and £9,594 and £25,744 per QALY gained respectively for non-T2DM patients.

In conclusion, the large variation around the ICERs when different random numbers and sampled patient profiles are used is of particular concern. In two different model runs of the ERG base-case, the ICER varied by as much as £7,000 per QALY gained. It is therefore the ERG's view that the company's model is of very limited value for the current decision problem and that results are to be interpreted with extreme caution.

2. BACKGROUND

In this section the ERG provides a review of the evidence submitted by Orexigen in support of naltrexone plus bupropion (NB32), trade name Mysimba® as a centrally acting anti-obesity product. We outline and critique the company's description of the underlying health problem and the overview of current service provision. The information is taken from Chapter 3 of the company submission (CS) with sections referenced as appropriate.

2.1 Critique of company's description of underlying health problem

The underlying health problem of this appraisal is overweight and obesity. According to Section 3.1 of the CS *"In clinical practice, body fatness is generally assessed by the BMI, calculated as body weight (kg) divided by height squared (m^2). The BMI range for normal weight is 18.5–24.9kg/m^2; overweight is 25–29.9kg/m^2; obese is 30–40kg/m^2 and morbidly obese is defined as >40kg/m^2".*[1]

In Section 3.4 of the CS the prevalence of overweight and obesity is reported *"Based on the 2014 Health Survey for England, a total of 11,126,000 adults (aged ≥16) were obese (BMI ≥30kg/m^2). In addition 15,825,000 adults are overweight[2] with around 30% or 4,747,500 having a BMI ≥27kg/m^2. Of these, an estimated 16% will have one or more weight-related comorbidity, equivalent to 779,680 patients. Therefore, a total of 11,905,680 adults in England are overweight or obese with one or more weight-related comorbidities".*[1]

In Section 3.1 of the CS it is noted that *"Men are more likely to be overweight; however women are more likely to be obese."* It is also noted that *"those aged 55–64 years are the most likely to be obese, while 16–24 year olds are least likely."*[1]

The CS states that **"***For both overweight and obesity, the fundamental cause is an energy imbalance between calories consumed from food and drink and calories expended through exercise and energy expenditure; over time, this imbalance results in abnormal or excessive fat accumulation*." The submission also highlights increased intake of foods that are high in fat and a decrease in physical activity levels as the most influential factor in increasing the prevalence of obesity. The CS also references a number of other factors influencing obesity.[1]

The CS describes how a number of health problems are associated with being overweight or obese and that the available literature focuses on those associated with obesity. They also state that *"because many people who are overweight will become obese in their lifetime, it is reasonable to assume the comorbidities listed are relevant to both populations".* These include T2DM, hypertension, heart disease, dyslipidaemia, coronary artery disease and stroke, respiratory effects, cancers, reproductive function and osteoarthritis.[1]

In Section 3.2 of the CS the company states *"Overweight and obesity also have a substantial mental health burden and can be associated with sleep apnoea and severe depression".*[1]

Section 3.4 of the CS states that *"In 2004, research by a House of Commons Select Committee estimated that 34,100 deaths were attributable to obesity. This equates to 6.8% of all deaths in England".*[3]

The economic burden of obesity is highlighted in the CS. *" A report from 2007 estimated that NHS costs attributed to elevated BMI were £4.2 billion, with indirect costs amounting to £15.8 billion.[4] This was expected to rise to £6.3 billion in 2015, £8.3 billion in 2025 and £9.7 billion in 2050.[3, 4]".*[1]

ERG comment: The ERG checked the references provided to support the statements in the submission. In general these were found to be appropriate. However the ERG noted a number of discrepancies:

- Although BMI measures of overweight and obesity cited in the CS match NICE guidelines,[5] the guidelines also emphasise that BMI should be interpreted with caution and that waist circumference in people with a BMI < 35kg/m2 should be considered. The guidelines also state that "*The use of lower BMI thresholds (23 kg/m2 to indicate increased risk and 27.5 kg/m2 to indicate high risk) to trigger action to reduce the risk of conditions such as type 2 diabetes, has been recommended for black African, African-Caribbean and Asian (South Asian and Chinese) groups*."[5]
- It was unclear how exactly numbers of adults who are overweight or obese with weight-related comorbidities in England quoted in the CS were derived. No source was cited for the estimated 16% with a weight-related comorbidity.
- The statement that women are more likely to be obese is incorrect. Twenty-seven percent of both genders are obese.[2] Women are more likely to be morbidly obese (BMI>40) than men (3.6% vs 2.2%) 68% of men were overweight or obese in 2015 compared to 58% of women.[3]
- Important variations for the prevalence of obesity have also been linked with social class. It has been suggested that this is associated with the degree of relative social inequality.[4]
- The studies supporting the link between excess weight and depression report an association only for those who are severely obese and/or have a chronic health condition.
- The report cited by the company on deaths associated with obesity referenced data collected in 2001.[3] According to the World Health Organisation, an estimated 9.6% of deaths among men and 11.5% of women are due to overweight and obesity in developed countries.[6] Applying these to England (2001 data) gives 52,500 not 34,100 deaths attributable to obesity as cited by the company.

2.2 Critique of company's overview of current service provision

The CS notes that in England *"Treatment is based upon a patient's BMI and what, if any, comorbidities are present, as outlined in Table 8"* (duplicated below).[1]

Table 2.1: Summary of treatment options for overweight and obese patients

BMI classification (kg/m2)	Waist circumference[a]			Comorbidities present
	Low	High	Very high	
Overweight (25-29.9)	1	2	2	3
Obesity I (30-34.9)	2	2	2	3
Obesity II (35-39.9)	3	3	3	4
Obesity III (40 or more)	4	4	4	4
Treatment options				
1	General advice on healthy weight and lifestyle			
2	Diet and physical activity			
3	Diet and physical activity, consider drugs			
4	Diet and physical activity, consider drugs; consider surgery			

Source: Table 8 of the CS[1]
Footnote: [a] for men, waist circumference of less than 94cm is low, 94–102cm is high and more than 102cm is very high. For women, waist circumference of less than 80cm is low, 80–88cm is high and more than 88cm is very high.
BMI = body mass index

The CS states that *"in NHS England, the initial standard of care is to advise lower-energy diets, increased physical activity and behavior modification. The exact nature of these treatments can vary in both style and intensity throughout NHS England and may be delivered by either dieticians, GPs or WeightWatchers®. For patients who have not achieved adequate weight loss (who have not reached their target weight loss, or who have reached a plateau) on such standard management, pharmacological treatment should be considered.*"[1]

In Section 3.3 of the CS it is stated that "*Currently in the EU, orlistat is the only available, orally effective, pharmacological product for weight management on the market; this is especially problematic given the complex aetiology of the disease across individuals (…)Due to its mechanism of action, orlistat is associated with several limitations, as detailed in Section 3.6. Therefore, the potential benefits of the addition of pharmacotherapy to standard management are not generally observed, as use of orlistat remains low."*[1]

In Section 3.3 of the CS it is reported that *"Surgery is only indicated for patients with a BMI ≥40kg/m^2 or between 35kg/m^2 and 40kg/m^2 with other significant disease, and who have failed all non-surgical measures, including intensive management in a Tier 3 service. Therefore, surgery should be considered a last resort for patients who have exhausted all other treatment options as seen by the limited number of surgeries conducted each year, and is therefore not considered an appropriate comparator to NB32, in line with the final scope for this submission."*[1]

The company states that *"NB32 can be used as an alternative first-line pharmacological treatment in patients for whom orlistat is contraindicated or is not utilized due to physician / patient choice, and patients who persevere with standard management despite the expected lack of effectiveness. NB32 should also be considered for patients who have not achieved adequate weight loss with orlistat treatment, or who did not comply with dietary requirements associated with orlistat, or were unable to tolerate orlistat treatment and who would otherwise revisit standard management measures."*[1]

ERG comment: The company provides an appropriate overview of the current provision of services in relation to overweight and obesity. However the following should be noted:

- Although the limitations of orlistat in terms of gastrointestinal adverse effects are appropriately highlighted, the CS does not provide data to support that the use of orlistat remains low. In England in 2014, pharmacies dispensed just over half a million items for treating obesity with a net ingredient cost of £15.3 million. All of these prescriptions were for orlistat.[3]
- Surgery provides better long-term outcomes for the morbidly obese (BMI>40).[5] A total of 6,032 bariatric surgery procedures (1,444 in male and 4,588 in women) were carried after a diagnosis of obesity in the year 2014-2015.[3]
- The ERG notes that NB32 is placed at first line in the clinical pathway as an alternative to orlistat and at second line in the pathway for those who have previously taken orlistat unsuccessfully. However in none of the main trials have patients previously taken orlistat.

3. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 3.1: The decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
Population	Adults who have a BMI of: ≥30kg/m^2 (obese) or ≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities	Adults who have a BMI of: ≥30kg/m^2 (obese) or ≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities	-	In line with the scope of the decision problem.
Intervention	Naltrexone-bupropion prolonged-release	Naltrexone-bupropion prolonged-release	-	In line with the scope of the decision problem. Note also that, in fact the intervention is an add-on to standard management.
Comparator (s)	Standard management without naltrexone-bupropion Orlistat (prescription dose)	Standard management without naltrexone-bupropion Orlistat (prescription dose)	-	In line with the scope of the decision problem. However, it is not clear what is meant by "Standard management without naltrexone-bupropion".
Outcomes	BMI Weight loss Percentage body fat Waist circumference Incidence of Type 2 diabetes Cardiovascular events Mortality Adverse effects of treatment Health-related quality of life	Weight loss Percentage body fat Waist circumference Incidence of Type 2 diabetes Cardiometabolic parameters Mortality Adverse effects of treatment Health-related quality of life	Key outcomes captured in pivotal trial programme	BMI and percentage body fat are not reported in the CS. The data on cardiovascular events are limited.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
Economic analysis	The cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year The time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared Costs will be considered from an NHS and Personal Social Services perspective	The cost effectiveness of treatments is expressed in terms of incremental cost per quality-adjusted life year The time horizon for estimating clinical and cost effectiveness reflects the lifetime of patients Costs are considered from an NHS and Personal Social Services perspective	-	In line with the scope of the decision problem.
Subgroups to be considered	People with Type 2 diabetes	People with Type 2 diabetes; the COR-DM study provides data for this subgroup	-	In line with the scope of the decision problem.
Special considerations including issues related to equity or equality	None specified	None specified	-	-

Source: CS[1]

BMI = body mass index; CV = cardiovascular; NB32 = naltrexone 32mg plus bupropion; NICE = National Institute for Health and Care Excellence; T2DM = type 2 diabetes mellitus

3.1 Population

The population is described in the scope as "Adults who have a BMI of:

- ≥ 30kg/m2 (obese) or
- ≥ 27kg/m2 to < 30kg/m2 (overweight) in the presence of one or more weight-related co-morbidities."[7]

The population in the Company Submission (CS) matches the scope.

3.2 Intervention

The intervention is described in the scope as naltrexone-bupropion prolonged-release. This is the same in the CS.

The indication for naltrexone-bupropion prolonged-release (32mg daily) or NB32 (UK brand name: Mysimba) is as follows:

"Mysimba is indicated, as an adjunct to a reduced-calorie diet and increased physical activity, for the management of weight in adult patients (≥18 years) with an initial Body Mass Index (BMI) of

- *≥30kg/m^2 (obese), or*
- *≥27kg/m^2 to <30kg/m^2 (overweight) in the presence of one or more weight-related co-morbidities (e.g., Type 2 diabetes, dyslipidaemia, or controlled hypertension)"*[1]

Treatment with Mysimba should be discontinued after 16 weeks if patients have not lost at least 5% of their initial body weight.[1] In most trials NB32 is continued throughout the trial, usually one year duration. The company states that "*For patients continuing treatment post 16 weeks, treatment should be continued as long as clinical benefit is observed*."[1] It is unclear how long treatment duration would be in clinical practice.

Mysimba is orally administered. Each tablet contains 8mg naltrexone and 90mg bupropion hydrochloride. Dose should be escalated for the first four weeks as follows:

- Week 1: One tablet in the morning
- Week 2: One tablet in the morning and one tablet in the evening
- Week 3: Two tablets in the morning and one tablet in the evening
- Week 4 and onwards: Two tablets in the morning and two tablets in the evening

In Section 2.3 Table 6 of the CS there is a statement "*Retreatment with NB32 is not routinely anticipated and thus not modelled.*"[1] The company was asked to justify why patients would not be retreated with naltrexone-bupropion for any subsequent weight gain after a successful treatment with the drug.[8] The company replied "*There are no data to indicate the effectiveness of retreatment with NB32 following successful treatment with NB32 and subsequent discontinuation and weight regain. If NICE thinks this is likely to happen in practice, an option for NICE is to consider that the current cost-effectiveness model assumes the same analysis for patients independent of whether they have received previous NB32 or not. Clinical rationale can inform the likelihood of retreatment success until evidence merges.*"[9]

3.3 Comparators

The comparators described in the scope are 'Standard management without naltrexone-bupropion' and 'Orlistat (prescription dose)'. These are the same in the submission.

However, the NICE scope does not specify what is meant by 'Standard management without naltrexone-bupropion'.

According to the CS, standard management consisted of customary diet and behaviour modification in three of the four main trials (COR-I, COR-II and COR-DM; CS, page 16-17).[1] In these three trials at baseline, weeks 12, 24, 26 and 49 (4, 16, 28 and 40 for COR-DM) patients received instructions to follow a hypocaloric diet (500 kcal/day deficit) and increase physical activity, and written behaviour modification advice.

In response to the ERG, the company stated that in the COR-I and COR-II studies "*Patients were encouraged to increase physical activity, with a prescription for walking starting with at least 10 minutes on most days of the week, and increasing this gradually to 30 minutes on most days of the week throughout the study. They were encouraged to lose weight and maintain weight loss, and were encouraged to follow the prescribed programme (as described). Participation in any other weight loss programme was not permitted. The use of meal replacements (such as Slim Fast® or Weight Watchers®) was discouraged, but occasional use did not necessitate withdrawal from the study. The prescribed exercise could be performed in a gymnasium or health club*."[9]

In COR-DM "*Patients were encouraged to increase physical activity, with a prescription for walking at least 30 minutes three times per week. Patients were encouraged to follow the prescribed programme. Participation in any other weight loss programme was not permitted. The use of meal replacements (such as Slim Fast® or Weight Watchers®) was discouraged, but occasional use despite contrary instructions did not necessitate withdrawal from the study. The prescribed exercise could be performed in a gymnasium."*[9]

In COR-BMOD standard management consisted of intensive behaviour modification. According to information provided by the company, it included "*three components: dietary instruction, closed group sessions, and prescribed exercise*".

BMOD consisted of group meetings lasting 90 minutes weekly for the first 16 weeks, every other week for the next 12 weeks and monthly thereafter. They included instructions to consume a balanced deficit diet and to increase to 180 min/week of planned, moderately vigorous, physical activity (CS, page 57).

In the COR-I and COR-II trials participants were not permitted to engage in weight loss programmes other than the prescribed programme of diet modification and exercise advice. This represents a more minimal approach to standard management than might be expected in practice. The COR-BMOD trial could be seen as best practice for standard management in that a more intensive intervention was delivered. Group sessions were included as well as dietary instruction and prescribed exercise. The choice of standard management has implications for the effectiveness and cost effectiveness of NB32 and these will be highlighted in this report.

The marketing authorisation for orlistat (UK brand name: Xenical) states that "The Committee for Medicinal Products for Human Use (CHMP) decided that Xenical's benefits are greater than its risks in conjunction with a mildly hypocalorific diet for the treatment of obese patients with a BMI greater or equal to 30 kg/m2, or overweight patients (BMI ≥28 kg/m2) with associated risk factors. The Committee recommended that Xenical be given marketing authorisation."[10] Orlistat comes as a capsule (120mg) to be taken three times a day.

The marketing authorisation further states that: "Xenical is given as one capsule taken with water just before, during, or up to one hour after each main meal. If a meal is missed or contains no fat, Xenical should not be taken. The patient should be on a diet in which about 30% of the calories come from fat,

and which is rich in fruit and vegetables. The food in the diet should be spread over three main meals. Treatment with Xenical should be stopped after 12 weeks if patients have been unable to lose at least 5% of their body weight since the start of treatment."[10]

In response to the draft scope, the Royal College of Physicians (RCP) pointed out that "The comparators seem reasonable but there is no direct head to head comparison with orlistat. Most of the trials with orlistat were conducted over 20 years ago, so caution should be exercised when making indirect comparisons; this is particularly true for conditions such as diabetes where background standard therapy (for glucose and lipids especially) may be very different now."[11] In addition, the RCP stated that "There is very little data on some ethnic groups (particularly people from Asia) within the trials with Naltrexone-Bupropion, so it may not be possible to make any firm conclusions for that group."[11]

3.4 Outcomes

None of the NB32 trials report BMI. In the CS this is explained as follows (CS, page 51):

"*Of note, change in BMI was not a pre-defined endpoint. Although this is an adequate research tool, it is limited in the assessment of an individual, as it does not consider different body morphologies (e.g. muscle vs adipose) and may be skewed by very high muscle mass.[12] In addition, some population groups, such as people of Asian family origin and older people, have comorbidity risk factors that are of concern at different BMIs (lower for adults of an Asian family origin and higher for older people)*.[5] *Therefore, alternative methods to measure body fatness, such as waist circumference, were utilised in the trials.*"[1]

However, NICE Clinical Guideline (NICE CG189, 2014[5]) states that:

- BMI should be used as a practical estimate of adiposity in adults
    - BMI should be interpreted with caution; waist circumference may be used in addition for patients with BMI <35kg/m2
    - Bioimpedance should not be used
- BMI should be interpreted with caution in muscular adults
    - Other populations, such as Asians and older patients, have comorbidity risk factors that are of concern at different BMIs
- Assessment of health risks associated with being overweight or obese should be based on BMI and waist circumference

Furthermore, BMI could easily have been calculated from data available in the trials.

In addition, 'cardiovascular events' are not reported in the CS. Instead the CS reports 'cardiometabolic parameters'. The FDA requested a trial: NB-CVOT to examine the risk of cardiovascular events, but this trial was terminated early. Where cardiovascular events are reported in the CSRs we will add them to this report.

3.5 Other relevant factors

No special considerations, including issues related to equity or equality, were specified (CS, page 14). A patient access scheme is not mentioned in the submission.

4. CLINICAL EFFECTIVENESS

The company conducted a systematic review to identify studies of NB32 and potential comparator therapies to treat adults who are overweight or obese. In Section 4.1 we critique this review.

4.1 Critique of the methods of review(s)

The systematic review conducted by the company formed the source of studies for both the NB32 direct evidence and the indirect treatment comparison between NB32 and orlistat. It was used to inform both efficacy and adverse event data.

4.1.1 Searches

The following paragraphs contain summaries and critiques of all searches related to clinical effectiveness presented in the company submission. The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[13] The submission was checked against the Single Technology Appraisal (STA) specification for company/sponsor submission of evidence.[14] The ERG has presented only the major limitations of each search strategy in the report.

The company submission stated that systematic review searches were undertaken in May 2016. Search strategies were reported in Appendix 2 of the CS for the following databases: Embase, MEDLINE, MEDLINE in-Process, Cochrane's CENTRAL, DARE and CDSR.

Additional searches of the following conference proceedings were reported for the last two years: International Congress on Obesity (ICO), European Congress on Obesity by the European Association for the Study of Obesity (ECO), American Diabetes Association (ADA), International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual European Congress and ISPOR Annual International Congress. In their response to clarification the company confirmed that the conference searches were conducted in June 2016 and provided the search terms.[9]

The CS also reported that the reference lists of existing systematic literature reviews and meta-analyses were checked for additional studies not identified by the main searches.

Searches utilised study design filters based on the Scottish Intercollegiate Guidelines Network (SIGN) Embase filters for RCTs, Observational Studies and Systematic Reviews.[15]

ERG comment:

- The database searches were clearly structured and documented. No language limits were applied.
- In their response to clarification the company confirmed that they searched Embase and MEDLINE simultaneously using a single database provider (Embase.com) and search strategy. This approach has limitations when using subject heading terms which could affect recall of results. Embase subject heading terms (Emtree) were used in the search strategy, and although simultaneous searching of Embase.com should automatically identify and search for equivalent MEDLINE subject heading terms (MeSH), it is not clear if this is the case for all potentially useful MeSH terms. Given the potential limitations of this approach, the ERG considered it preferable to search each database separately, or at least to ensure inclusion of both Emtree and MeSH terms in the search strategy.
- The ERG noted the use of study design filters in the Cochrane Library searches of CDSR, DARE and CENTRAL. It was considered that this was an overly restrictive approach given that these resources are already filtered by study design. Of particular concern was the search

of CENTRAL, which when rerun by the ERG yielded approximately 65 additional results without the study design filters. The ERG requested that the company rerun this search and screen these additional papers to confirm that no relevant papers had been missed. In their response to clarification the company responded "*Searches were conducted again by applying the CENTRAL limit in the Cochrane Library instead of using the study design filters, as was done originally. This found only five additional unique papers from which three were deemed relevant. However, these three potentially relevant studies were published after June 2016, when the original searches were conducted. As such, no additional studies were included from this approach*."[9]

- Section 4.10.1 stated "The search strategy used to identify RCT evidence for NB32 and orlistat 120mg TID is described in Section 4.1."[1], therefore the same limitations as described above will have applied.
- No mention was made in Section 4.12 of the company submission with regard to how adverse events data were identified. The ERG queried this omission and asked for confirmation that the results of searches reported in Appendix 2 of the CS were screened for adverse events. Guidance by the Centre for Reviews and Dissemination (CRD)[16] recommends that if searches have been limited by a study design filter, additional searches should be undertaken to ensure that adverse events that are long-term, rare or unanticipated are not missed. The company responded: "*No additional searches to those reported in Section 4.1 and Appendix 2 were conducted to identify adverse event (AE) data, but results retrieved were screened for AEs*." This issue is further discussed in Section 4.1.2 of this report.

4.1.2 Inclusion criteria

The eligibility criteria used in the search strategy of the review for randomised controlled trials (RCTs) and non-RCTs is presented in Table 4.1.

Table 4.1: Eligibility criteria for trials to be included in the systematic review

Criteria	Inclusion criteria	Exclusion criteria
Population	Adults who are obese (BMI ≥30kg/m2) or overweight, according to one of the following definitions: • 25kg/m2 to 29.9kg/m2 • ≥27kg/m2 to <30kg/m2 • >28kg/m2 with one or more weight-related comorbidity (T2DM, dyslipidaemia and/or controlled hypertension)	Healthy volunteers Children (age <18 years) Diseases other than that specified in inclusion criteria
Study design	RCTs Non-RCTs Systematic reviews and meta-analyses of RCTs[a]	*In vitro* studies Preclinical studies Comments, letters, editorials Case reports, case series Non-systematic reviews Observational studies
Intervention	Studies assessing at least one of the following interventions will be included: Naltrexone-bupropion Orlistat	Studies that do not assess at least one of the included interventions will be excluded

Criteria	Inclusion criteria	Exclusion criteria
Comparator	Comparator therapies may include one of the following: Behavioural interventions Lifestyle or dietary modifications Any treatment listed under the interventions Any other pharmacological treatments for obesity or weight management	Studies will not be excluded on comparator therapy if it includes at least one of the treatments listed under the interventions
Study duration	All trials with total randomised phase duration >1 year are included	Studies with <1-year duration
Language	Studies published in English were included Studies published in non-English languages were flagged	Studies will not be excluded on the basis of publication language
Source: Table 10 of the CS[1] Footnote: a, Systematic reviews and meta-analyses of RCTs were identified and flagged. Bibliographies of these systematic reviews will be screened to check if literature searches have missed any potentially relevant studies. BMI = body mass index; RCT = randomised controlled trial; T2DM = Type 2 diabetes mellitus.		

ERG comment:

- Although non-RCT studies were eligible for inclusion, they were not considered further once sufficient RCTs were found. The company was asked to clarify the exclusion of non-RCTs. They responded that *"Both RCTs (randomised controlled trials) and non-RCTs were identified through SLR (systematic literature review), and screened for AEs. However, non-RCT evidence was not formally considered as part of comparative safety assessments as RCT data were available for the intervention and comparators of interest to the decision problem. This included longer-term safety data to that available from the pivotal trial programme*." Although this may be acceptable for effectiveness data, it is not normally acceptable for adverse events. Non-RCT studies should have been assessed for long-term follow-up and reporting of rare adverse events.[16] Additionally, bibliographic details of the nine non-RCTs should have been provided. However, in the case of this technology assessment, the ERG did not find any relevant non-RCT studies of NB32 that were missed or inappropriately excluded.
- The inclusion criteria state that "*Studies published in non-English languages were flagged*."[1] The company was asked to clarify the methods for dealing with these studies and responded "*Non-English language studies were to be included if sufficient evidence from English language articles was not available. In light of the completeness of English language RCTs, all non-English language studies were excluded*."[9] The ERG noted that 44 full text articles were excluded but a complete list of these articles was not provided so it was not possible to ascertain if any relevant non-English language studies had been excluded.
- Studies that do not assess at least one of the included interventions (naltrexone-bupropion or orlistat) were excluded. This means that studies comparing a behavioural intervention with placebo (or waiting list control) have been excluded as were studies comparing different types of behavioural interventions. According to the scope, these studies should have been included; this would have allowed an indirect comparison of naltrexone-bupropion versus different types of behavioural interventions.

4.1.3 Critique of data extraction

The company stated that "*All relevant data were extracted from the included full text of articles by one reviewer and quality checked against the original source by a second reviewer*".[1]

ERG comment: Although the company stated that two reviewers were involved in the data extraction of included studies, it was unclear how discrepancies were resolved (e.g. use of a third reviewer). Although it is good practice to include this detail when reporting a systematic review, we believe that overall the data extraction was carried out appropriately.

4.1.4 Quality assessment

The CS stated that quality assessment of included studies was done "in accordance with the NICE-recommended checklist for RCT assessment of bias".[1] Elements assessed were randomisation, allocation concealment, comparability of groups, blinding of care providers, patients and outcome assessors and drop out, selective reporting of outcomes and use of intention to treat analysis and appropriate methods for dealing with missing data.

ERG comment: Study quality appeared to have been assessed using appropriate tools.

4.1.5 Evidence synthesis

Two types of evidence synthesis are described in the CS: a meta-analysis of the NB32 trials and an indirect comparison comparing NB32 with orlistat.

The meta-analysis

To compare and pool the relative treatment effects between the four trials comparing NB32 and placebo (COR-I, COR-II, COR-BMOD and COR-DM), a frequentist pairwise meta-analysis was performed to assess the following outcomes:

- At least a 5% reduction in weight at one year from baseline (the one year time point ranged from 52 to 57 weeks). This was a dichotomous outcome.
- Mean percentage weight change from baseline at one year (the one year time point ranged from 52 to 57 weeks). This was a continuous outcome.

The ERG asked for clarification as to why these outcomes had been selected for the meta-analysis. The company responded that "*The outcome of 5% reduction in weight from baseline was incorporated as per the European Medicines Agency (EMA) licence and associated treatment stopping rules; whereas the mean % weight change from baseline was incorporated to account for the overarching treatment effect of each regimen. Meta-analysed results for alternate outcomes were not required for the de novo model, and were therefore not produced*."[9]

The NB-CVOT study was excluded from all meta-analyses, due to the trial design, objective, and patient population, being different from the other studies.

The frequentist pairwise meta-analysis was performed using R (version 3.3.1) using the metafor package.[17, 18] The pairwise meta-analysis, presents relative treatment effects per trial, and an overall 'pooled' relative treatment effect for placebo versus NB32 which was calculated using a random effects model.[19] To further evaluate the trial-heterogeneity, sensitivity analyses were performed for the three non-T2DM trials, and for the non-T2DM trials excluding the COR-BMOD trial, as patients received intensive behaviour modification. The statistical heterogeneity of the pairwise meta-analysis was assessed using I^2, where the I^2 value describes the percentage of total variation across studies that is due to heterogeneity rather than chance.[20] The mITT populations were used in the meta-analyses. Results for the number with ≥ 5% reduction in weight (binary outcome) were reported as odds ratios (OR) and

results for the mean percentage change in weight from baseline (continuous outcome) were reported as mean differences (MD) both with 95% confidence intervals (CI).

ERG comment: The meta-analyses used appropriate statistical methods. Only two outcomes were included in the meta-analysis and both were measures of weight loss, these were also the two co-primary outcomes in the COR trials. The company stated that other outcomes were not meta-analysed as they were not needed for the economic model. This seems to be reasonable as the other outcomes were reported for the COR trials, and given the heterogeneity in terms of populations and background therapy (see below) additional meta-analyses may not have been appropriate.

Subgroup analyses were performed to explore the heterogeneity by splitting the studies into those containing only type 2 DM (T2DM) patients and those excluding type 2 DM patients. The standard management received in the COR-BMOD trial was more intensive than in the COR-I and II trials. The CS states that "*these differences (the presence or absence of T2DM and the intensity of the diet and exercise programme) between the trial designs are likely to explain the heterogeneity in results between the four trials*" (CS, Section 4.91, page 111). The ERG agrees that there was clinical and statistical heterogeneity between the four COR trials and that because of this the results from the separate analyses for T2DM and no T2DM should be used.

The indirect comparison

An indirect treatment comparison (ITC) was performed to compare NB32 with orlistat (120mg TID), using placebo as a common comparator.

ITC were performed to compare NB32 and orlistat for the following outcomes:

- Mean percentage weight change from baseline at one year (the one year time point ranged from 52 to 57 weeks [continuous outcome])
- At least 5% reduction in weight at one year from baseline (the one year time point ranged from 52 to 57 weeks [dichotomous outcome])

Some data imputations were required to maximise inclusion of evidence in the analyses, and the methods of imputation are described in Appendix 10 of the CS. The analysis used the mITT populations.

Odds Ratios (OR) were used as the effect size for ≥ 5% reduction in weight and mean differences (MD) for the mean percentage change in weight from baseline.

To investigate the effect of T2DM and to populate the economic model (in which results from the ITC were applied according to individual patient T2DM status), if data were available then all the analyses and sensitivity analyses were performed separately for:

- Trials where T2DM is part of the trial inclusion criteria (T2DM analysis)
- Trials where T2DM is part of the trial exclusion criteria (no T2DM analysis)
- All trials regardless of T2DM (any T2DM analysis)

To assess the effects of weight loss in trials where a large proportion of patients had comorbidities, a sensitivity analysis was performed excluding trials where ≥75% of patients ≥1 comorbidity (hypertension, dyslipidaemia, or T2DM). Due to anticipated heterogeneity with respect to the duration of and therapies received during the lead-in periods, sensitivity analyses were also performed excluding those trials incorporating lead-in periods.

The specific type and intensity of standard management varied between the trials, although treatment arms within the same trial received the same standard management. For the analysis, it was therefore

assumed that the additional treatment benefit from the standard management was additive but that the relative treatment effect between treatment arms would be unaffected. Further sensitivity analyses were performed excluding studies with 'intensive' behaviour modification, and excluding trials with lead-in periods or 'intensive' behaviour modification.

A Bayesian NMA was performed for each outcome using the available data (CS, Table 31 and Table 32). Markov Chain Monte Carlo (MCMC) methods were used which combine prior distributions with the data to construct a posterior distribution of parameters of interest upon which to base summary results. All models were fitted using WinBUGS (version 14),[21] via R (version 3.3.1).[17] An initial 50,000 iterations were discarded as the 'burn-in' period, which was assessed by running two chains using different starting values and assessing convergence using Brooks-Gelman-Rubin plots.[22] Then, 10,000 samples (posterior distribution) were used for obtaining summary estimates. In total, 10,000 samples were deemed sufficient for each of the different analyses as the Monte Carlo error was less than 5% of the standard deviation.[23] Therefore, the samples could be used directly in the economic model, preserving the correlation between treatment effects and avoiding the need to make assumptions regarding the shape of the posterior distribution. Autocorrelation was assessed to determine whether samples were highly correlated, a thinning interval of five was applied to ensure that the chain was mixing well and was representative of the posterior distribution. The goodness-of-fit was assessed using the total residual deviance and tested using a chi-squared test. Random effects and fixed effect models were used; however, random effect results are only presented for the 'any T2DM' analysis. Random effects results are not presented for the T2DM only and non-T2DM analyses, as the models failed to update effectively using the recommended priors, likely due to the low number of studies. The models and prior distributions used for the two outcomes were those described in the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 2.[24]

ERG comment: The ERG re-ran the Bayesian NMA for both the binary and continuous weight loss outcomes and reproduced the results reported in the CS for the three analysis groups: T2DM, no T2DM and any T2DM patients. The modelling used the code supplied in the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 2[24] and the analysis methods were appropriate. Model fit was tested and the results were reported. The decision to present only fixed effect model results for the T2DM and no T2DM subgroups was correct as the ERG also found that there were problems with model convergence for these models, especially for the T2DM analyses. The fixed effect model provided the best fit to the data and results that are likely to be more reliable. There was no need to evaluate inconsistency in the analyses as they were straightforward indirect comparisons between NB32 and orlistat using placebo as the common comparator. Appropriate sensitivity and subgroup analyses were used to explore differences resulting from the inclusion or exclusion of patients with T2DM and those trials using intensive behaviour modification as background therapy.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

4.2.1 Overview of the evidence in the submission

The company identified 36 relevant RCTs. The CS stated that "*Of the 36 included RCTs, 5 studies investigated treatment with NB32…., while the remaining 31 studies investigated treatment with orlistat*."[1] The five studies of NB32 will be discussed in this section and are listed in Table 4.2. The studies of orlistat were used to form an indirect comparison with NB32 and will be discussed in Section 4.4 of this report.

Table 4.2: List of relevant RCTs

Trial name	Population	Intervention[a]	Comparator
COR-I[25]	Adults with uncomplicated obesity or who were overweight with dyslipidaemia or hypertension	Naltrexone 32 mg per day + bupropion 360 mg per day (NB32) OR Naltrexone 16 mg per day + bupropion 360 mg per day (NB16)	Placebo
COR-BMOD[26]	Adults with uncomplicated obesity or who were overweight with dyslipidaemia or hypertension	Naltrexone 32 mg per day + bupropion 360 mg per day (NB32) + BMOD	Placebo + BMOD
COR-II[27]	Adults with uncomplicated obesity or who were overweight with dyslipidaemia or hypertension	Naltrexone 32 mg per day + bupropion 360 mg per day (NB32)	Placebo
COR-DM[28]	Adults with T2DM and BMI ≥ 27 and ≤ 45 kg/m^2	Naltrexone 32 mg per day + bupropion 360 mg per day (NB32)	Placebo
NB-CVOT[29]	Adults with a BMI of 27 to 50 and who had characteristics associated with an increased risk of CV outcomes [b]	Naltrexone 32 mg per day + bupropion 360 mg per day (NB32)	Placebo

Source: Tables 11 and 12 and text of section 4.2 of the CS[1]
Footnote: a) Two tablets of NB32 or placebo were taken twice a day (each tablet contained 8mg naltrexone hydrochloride and 90mg bupropion hydrochloride) b) terminated early (after 50% interim analysis)
BMI = body mass index; BMOD = intensive behaviour modification; COR = Contrave obesity research; CV = cardiovascular; CVOT = cardiovascular outcomes trial; DM = diabetes mellitus; RCT = randomised controlled trial; T2DM = type 2 diabetes mellitus

A further trial, IGNITE, unpublished at the time of the systematic review, was identified by the company and presented as supporting evidence.[30]

The company stated that there were no relevant ongoing trials. However in the background section of the CS two trials were mentioned "*a further Phase IV study to assess the effect of NB32 on the occurrence of MACE in overweight and obese patients was requested. Data from this trial are due in 2022. The CHMP also requested additional assessment of the pharmacokinetics of NB32 in patients with renal impairment and in patients with hepatic impairment, as the submitted trials did not collect such data, nor did the Phase III programme allow a direct evaluation of safety in these patient groups. Such a trial is ongoing*."[1]

The company was asked to provide details of these studies and to indicate if any interim data were available.[8] The company replied regarding the MACE study that "*Study synopsis is provided as an attachment. No information related to the new MACE study has been published or is available on any bibliographic database as it is currently still in the planning stage*".[9] The study is a multicentre,

randomised, double-blind, placebo-controlled study of the effect of NB32 on the occurrence of major adverse cardiovascular events (MACE) in overweight and obese adults with cardiovascular disease. Based in the US, the trial will aim to enrol 8,000 patients. It will have a lead-in period of two weeks, and a treatment period estimated to last for up to six years until the targeted number of adjudicated MACE events (378) has been reached. The primary MACE composite comprises the first occurrence of CV death, nonfatal myocardial infarction (MI), and nonfatal stroke.[31]

Regarding the trial of patients with renal impairment and patients with hepatic impairment, the company replied "*Study synopsis are provided as an attachment. As both the renal and hepatic impairment studies are small phase I studies requested by regulatory agencies, no information related to these studies have been published or made available on clinical study databases, such as clinicaltrials.gov.*"[9] Both studies aimed to enrol 32 to 48 participants. One was "*to evaluate the effect of hepatic impairment on the PK of naltrexone, bupropion, and their major active metabolites following a single oral dose of NB in subjects with varying degrees of hepatic function*."[32] And the other was "*To evaluate the effect of renal impairment on the PK of naltrexone, bupropion, and their major active metabolites following a single oral dose of NB (total dose of 16 mg naltrexone and 180 mg bupropion) in subjects with varying degrees of renal function*."[33]

The company stated that "*This submission focuses on data from the four pivotal RCTs: COR-I, COR-II, COR-BMOD, and COR-DM with only longer-term efficacy and safety data used to predict maintenance of pivotal trial outcomes presented from the NB-CVOT study and supported with data from the IGNITE study.*"[1] Accordingly, the four pivotal RCTs: COR-I, COR-II, COR-BMOD, and COR-DM will be discussed in some detail in Section 4.2.2 of this report whilst NB-CVOT and IGNITE will be discussed more briefly in Section 4.2.7.

All trials included patients who were obese or overweight with comorbidities. COR-I, COR-II and COR-BMOD excluded patients with diabetes but in COR-DM all patients had type two diabetes.

None of the trials compared NB to orlistat, a comparator specified in the NICE scope.[7] All the main trials compared NB32 to placebo. COR-I also included a treatment arm where patients received NB16.[25]

In both arms of the trials patients received customary diet and behaviour modification. According to the CS "*This included a hypocaloric diet (500 kilocalorie [kcal] per day deficit based on the World Health Organization [WHO] algorithm for calculating resting metabolic rate) as well as instructions on increasing physical activity (COR-I and COR-II), or more intensive behaviour modification counselling (COR-BMOD)*." This represents '*standard management without naltrexone-buproprion*' as specified in the NICE scope.[7] More detail is provided in Section 3.3 of this report.

ERG comment:

- The CS appropriately focuses on the four main NB32 RCTs. However these all compare NB32 to placebo with both arms receiving standard care. The ERG draws to the attention of the committee that no trials directly compare NB32 to orlistat as specified in the NICE scope.[7]
- The ERG also notes that standard care varies between the trials in that COR-BMOD has a more intensive form of behavioural management.
- The ERG confirms that evidence from the ongoing trials could not have been incorporated into the CS. However the ERG draws the attention of the committee to the ongoing MACE trial.[31]
- The ongoing investigations into patients with renal or hepatic impairment are drawn to the attention of the committee. Currently as stated in the CS, "*Patients with end-stage renal failure*

or severe renal or hepatic impairment are listed as a contraindicated patient population in the Mysimba SmPC."[1]

4.2.2 Overview of the direct evidence

This section focuses on the four main trials: COR-I, COR-II, COR-BMOD and COR-DM. Further details of their design can be found in Table 4.3.

Table 4.3: Trial designs of included NB32 studies

Trial name	Location	Number of participants	Trial design and duration	Primary outcome
COR-I[25]	34 study sites in the US	1,742	Phase III, multicentre, randomised, double-blind placebo-controlled 56 week study	Percentage of change in total body weight and proportion of patients with ≥ 5% decrease in total body weight at week 56
COR-BMOD[26]	9 study sites in the US	793	Phase III, multicentre, randomised, double-blind placebo-controlled 56 week study	Percentage of change in total body weight and proportion of patients with ≥ 5% decrease in total body weight at week 56
COR-II[27]	36 study sites in the US	1,496	Phase III, randomised, parallel-arm, double-blind, placebo-controlled, 56 week study	Percentage of change in total body weight and proportion of patients with ≥ 5% decrease in total body weight at week 28
COR-DM[28]	53 study sites in the US	505	Phase III, multicentre, randomised, double-blind placebo-controlled 56 week study	Percentage of change in total body weight and proportion of patients with ≥ 5% decrease in total body weight at week 56
Source: Table 12 of the CS[1] BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus; US = United States				

All trials were multicentre and all were conducted in the US. All had a joint primary outcome of percentage change in total body weight and proportion of patients with >5% decrease in total body weight. Three trials measured outcomes at week 56.[25, 26, 28] One trial, COR-II measured the primary outcome at 28 weeks. In COR-II, NB32 patients who had lost less than 5% of their body weight at visits between weeks 28 and 44 were re-randomised to continue with NB32 or escalate to NB48.

The four main trials included 4,536 patients. Of these 2,510 patients were randomised to NB32, 578 to NB16 (in COR-I) and 1,448 randomised to placebo.

ERG comment:

- As all of the trials were conducted in the US, participant characteristics may not reflect a UK population particularly in terms of ethnicity. Patient characteristics will be discussed later in this section.
- As all of the trials were conducted in the US, standard care may differ from a UK setting. Differences between the trials in terms of standard care have already been highlighted in Section 3.3 of this report.
- It is also possible that standard care varied within the trials given the number of centres (34 centres for COR-I, 36 for COR-II, nine for COR-BMOD and 53 for COR-DM).
- Three trials measure the primary outcome at 56 weeks. Although this is acceptable in terms of weight loss, there is no information on maintenance of weight loss after this time. The CS states that "*For patients continuing treatment post 16 weeks, treatment should be continued as long as clinical benefit is observed*."[1] It is unclear how long patients would continue to take the drug in practice.
- The licensing for NB32 indicates that it should be discontinued after 16 weeks if patients have not lost at least 5% of their initial body weight. However the main trials do not incorporate this stopping rule as the licensing was based on evidence found in the trials. The company stated "*As a result of pooled, post-hoc analyses of the COR trials that showed a strong relationship between early weight loss and clinically meaningful longer term weight loss, the license terms for NB32 include more prescriptive discontinuation rules*."[9]
- The ERG notes that the primary outcome includes ≥ 5% decrease in total body weight. The CHMP recommended investigation of ≥10% and this outcome is presented in the submission for individual trials but not for the meta-analyses. The results section of this report will also present results for patients with ≥ 10% weight loss.

Participant inclusion and exclusion criteria are shown in Table 4.4.

Table 4.4: Participant inclusion and exclusion criteria in the NB32 trials

Trial name	Patient age	Patient BMI	Includes patients with diabetes?	Main exclusion criteria relating to obesity
COR-I[25] COR-II[27] COR-BMOD[26]	18 to 65	BMI 30 to 45 kg/m2 and uncomplicated obesity OR BMI 27 to 45 kg/m2 and controlled hypertension and / or dyslipidaemia	No	Any anorectic or weight loss agents Participated in a weight loss management program concurrent to trial (COR-I and II) or within one month prior to randomisation(COR-BMOD) Weight change of > 4 kg within 3 months prior to randomisation Obesity of known endocrine origin History of surgical or device intervention for obesity History of treatment with, hypersensitivity or intolerance to bupropion or naltrexone
COR-DM[28]	18 to 70	BMI 27 to 45 kg/m2	Yes, all had T2DM	Type 1 diabetes Any anorectic or weight loss agents Obesity of unknown endocrine origin other than DM Loss or gain of > 5 kg within 3 months prior to screening Participated in a weight loss management program within one month prior to randomisation History of surgical or device intervention for obesity Treatment with bupropion or naltrexone within 12 months prior to screening

Source: Table 12 of the CS[1]
BMI = body mass index; BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus; T2DM = type 2 diabetes mellitus

Participant inclusion criteria for age and BMI are similar across the four main trials. As previously mentioned, one trial was conducted exclusively in patients with type 2 diabetes mellitus[28] whilst the other three excluded patients with diabetes. All trials included patients with a relatively stable weight and excluded obesity of endocrine origin. Other exclusions were patients were prior use of any anorectic or weight loss agents and those with a history of surgery or device intervention.

ERG comment:

- The ERG notes that evidence for diabetic patients was based on one trial of 505 participants.
- Inclusion and exclusion criteria appear to be reasonable for the main trials. The ERG draws to the attention of the committee that prior use of orlistat was an exclusion criterion in all four COR trials. Therefore the effect of NB32 on those who have failed on orlistat has not been examined.

Participant characteristics are displayed in Table 4.5.

Table 4.5: Participant characteristics in the NB32 trials

	COR-I[25]		COR-BMOD[26]		COR-II[27]		COR-DM[28]	
	NB32	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
Age, mean years (SD)	44.4 (11.1)	43.7 (11.1)	45.9 (10.4)	45.6 (11.4)	44.3 (11.2)	44.4 (11.4)	54.0 (9.1)	53.5 (9.8)
Age range (min, max)	19, 65	18, 66	19, 65	19, 64	18, 65	18, 65	20, 72	27, 70
Sex, female, n (%)	496 (85)	496 (85)	528 (89.3)	185 (91.6)	847 (84.6)	420 (84.8)	195 (58.2)	90 (52.9)
Ethnicity, n (%)								
White	440 (75)	440 (76)	405 (68.5)	149 (73.8)	835 (83.4)	414 (83.6)	261 (77.9)	140 (82.4)
Black	106 (18)	110 (19)	145 (24.5)	44 (21.8)	133 (13.3)	72 (14.5)	63 (18.8)	18 (10.6)
Asian	6 (1.0)	4 (0.7)	6 (1.0)	2 (1.0)	12 (1.2)	4 (0.8)	7 (2.1)	5 (2.9)
Other	31 (5.4)	27 (4.6)	35 (6.0)	7 (3.5)	21 (2.1)	5 (0.8)	4 (1.2)	7 (4.1)
BMI, mean kg/m^2 (SD)	36.1 (4.4)	36.2 (4.0)	36.3 (4.2)	37.0 (4.2)	36.2 (4.5)	36.1 (4.3)	36.4 (4.8)	36.4 (4.5)
Obesity class, n (%)								
BMI < 30 kg/m^2	18 (3.1)	5 (0.9)	8 (1.4)	1 (0.5)	25 (2.5)	14 (2.8)	18 (5.4)	11 (6.5)
BMI ≥30 and <35 kg/m^2	224 (38.4)	217 (37.3)	207 (35.0)	64 (31.7)	398 (39.8)	186 (37.6)	111 (33.1)	49 (28.8)
BMI ≥35 and <40 kg/m^2	204 (35.0)	229 (39.4)	230 (38.9)	79 (39.1)	316 (31.6)	191 (38.6)	110 (32.8)	64 (37.6)
BMI ≥40 kg/m^2	137 (23.5)	130 (22.4)	146 (24.7)	58 (28.7)	262 (26.2)	104 (21.0)	96 (28.7)	46 (27.1)
Other, n (%)								
Weight, mean kg (SD)	99.7 (15.9)	99.5 (14.3)	100.2 (15.4)	101.9 (15.0)	100.3 (16.6)	99.2 (15.9)	104.2 (18.9)	105.1 (17.0)
Smoker, n (%)	65 (11)	65 (11)	0*	0*	108 (10.8)	52 (10.5)	38 (11.3)	15 (8.8)
Hypertension, n (%)	130 (22)	113 (19)	86 (14.6)	37 (18.3)	212 (21.2)	106 (21.4)	212 (63.3)	103 (60.6)
Dyslipidaemia, n (%)	284 (49)	288 (50)	270 (45.7)	81 (40.1)	560 (55.9)	263 (53.1)	280 (83.6)	145 (85.3)
Alcohol use, n (%)	254 (43.6)	244 (42)	251 (42.5)	100 (49.5)	462 (46.2)	217 (43.8)	96 (28.7)	69 (40.6)
History of depression	66 (11.3)	73 (12.6)	83 (14.0)	31 (15.3)	131 (13.1)	76 (15.4)	29 (8.7)	14 (8.2)
History of anxiety	29 (5.0)	18 (3.1)	19 (3.2)	7 (3.5)	47 (4.7)	30 (6.1)	10 (3.0)	9 (5.3)
Statin use	11.5	8.6	9.1	8.4	11.7[$]	13.1	49.3	45.9

Source: Table 15 of the CS[1] and CSRs for COR-I[34], COR-BMOD[35], COR-II[36] and COR-DM[37]

Footnote: *Only non-smokers were eligible for the COR-BMOD trial. $ Includes NB48

BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus

Mean age was approximately 44 years apart from in COR-DM where participants were older (mean age 54)[28] The majority of participants were female although COR-DM had a more even distribution of female and male participants.[28] The majority of participants across the trials were of white ethnicity. Approximately 15% of participants were Black or African American. Just 1% were of Asian origin. Participants in the trials tended to be obese rather than overweight with an average BMI of 36 to 37. Approximately 2% had a BMI of < 30 (overweight). Hypertension was present in approximately 20% of patients across the COR trials although in COR-DM as expected over 60% had hypertension. Similarly, dyslipidaemia was present in approximately half of participants in the COR trials but in approximately 84% of the COR-DM patients.

ERG comment:

- Overweight patients in addition to obese patients were included in the NICE scope.[7] However there is only a very small percentage of patients who are overweight in the trials. Therefore the ERG draws to the attention of the committee that this population is not well represented.
- The majority of participants in the trials are female. The ERG draws to the attention of the committee that this does not reflect the distribution of obesity according to gender. Men in England are more likely to be overweight or obese (68% vs 58% in 2015).[3]
- The ERG draws to the attention of the committee that Asian patients are not well represented in the trials so results may not be applicable to these ethnic groups.

4.2.3 Direct evidence: Quality assessment

Table 4.6 presents the company's quality assessment of the four main trials with comments from the ERG.

Table 4.6: Quality assessment of included NB32 trials

Study question	Company's assessment of risk of bias				ERG comments
	COR-I[25]	COR-BMOD[26]	COR-II[27]	COR-DM[28]	
Was randomisation carried out appropriately?	Low	Low	Low	Low	Methods in all trials were appropriate.
Was the concealment of treatment allocation adequate	Low	Low	Low	Low	Methods in all trials were appropriate.
Were the groups similar at the outset of the study in terms of prognostic factors, for example severity of disease?	Low	Low	Low	Low	Methods in all trials were appropriate.
Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	Low	Low	Low	Low	Methods in all trials were appropriate.
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	Low	Low	Low	Low	Company noted that more patients dropped out of NB32 groups due to adverse effects.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	Low	Low	Low	Low	Methods in all trials were appropriate.
Did the analysis include an intention-to-treat-analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Low	Low	Low	Low	Analyses were performed on the modified ITT for all trials. ITT was included as a sensitivity analysis. All those in the ITT analysis had to have a post-baseline body weight measurement during the treatment phase and for the mITT the measurement was while on the drug.
Was statistical powering such to detect a significant difference between treatment groups?	Low	Low	Low	Low	Methods in all trials were appropriate.
Source: CS Appendix tables 3 to 6[38] BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus					

ERG comment:

- Apart from the mITT analyses, the ERG agrees that the four main trials were of high quality and attempts were made to lower the risk of bias. Methods of analysis relating to intention-to-treat will be discussed below. In terms of dropout, the ERG noted that more patients dropped out of NB32 groups due to adverse events. The ERG was concerned that higher rates of adverse events (especially nausea – see Section 4.2.5 of this report) in the intervention arm could have resulted in un-blinding of participants.

The main results presented in the CS were based on a modified intention-to-treat (mITT) analysis. According to the CS this was defined as "*all randomised patients with a post-baseline body weight measurement obtained while the patient remained on study medication*". Missing data were imputed using the LOCF method for primary analysis.[1]

Table 4.7 presents the numbers of patients randomised and the numbers included in the mITT analysis for each trial.

Table 4.7: Randomisation and analysis sets

	No randomised		**Company's 'Modified ITT analysis set'¶**	
Trial name	NB32	Pbo	NB32 n (% of randomised)	Pbo n (% of randomised)
COR-I[25]	583	581	471 (80.8)	511 (88.0)
COR-BMOD[26]	591	202	482 (81.6)	193 (95.5)
COR-II[27]	1001	495	825 (82.4)	456 (92.1)
COR-DM[28]	335	170	265 (79.1)	159 (93.5)
Source: Section 4.5 of the CS[1] and Appendix 5 of the CS[38] Footnote: ¶ All randomised patients with a post-baseline body weight measurement obtained while on study drug BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus				

It can be seen that using the modified ITT analysis loses approximately 19% of the patients randomised to NB32 and 9% of those allocated to placebo.

ERG comment:

- The main results presented in the company submission were based on a modified intention-to-treat (mITT) analysis. This analysis includes only those patients who have a baseline and at least one post-baseline measurement whilst on the study drug. Patients who discontinued without providing follow-up weight assessments were excluded. The use of the mITT population is likely to be biased as the reasons why a patient discontinued trial treatment or failed to return for post-baseline weight assessments could be related to the efficacy or safety of the drug. Patients who were not seeing a satisfactory weight loss or experiencing side effects are more likely to stop taking the study drug. Results for the true intention-to-treat analysis should be the main data presented in the submission as this includes all patients in the treatment arms to which they were originally randomised. In our report we present the ITT results in addition to the mITT results. The only study where full ITT results were not available was COR-BMOD.

4.2.4 Direct evidence: Efficacy results

The main results of the modified intention-to-treat analysis presented in the CS are shown in Table 4.8. Tables 4.9 and 4.10 compare the mITT results for the primary outcomes with the two methods of ITT analysis **(**weight regain imputation method and using baseline-carried forward analysis).

Table 4.8: Main results of NB32 trials (mITT analysis)

	COR-I[25]		**COR-II[27]***		**COR-BMOD[26]**		**COR-DM[28]**	
	NB32	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
N	471	511	825	456	482	193	265	159
Baseline weight, mean kg (SD)	100.2 (16.3)	99.3 (14.3)	100.7 (16.7)	99.3 (16.0)	100.7 (15.4)	101.9 (15.0)	104.2 (18.9)	105.0 (17.1)
End of study weight, mean kg (SD)	94.2 (17.4)	98.0 (15.2)	94.2 (17.6)	97.2 (16.2)	91 (17.1)	96.4 (17.1)	101.0 (19.7)	103.0 (17.3)
Percent change from baseline at end of study, LS mean (SE)	-6.1 (0.3)	-1.3 (0.3)	-6.5 (0.2)	-1.9 (0.3)	-9.3 (0.4)	-5.1 (0.6)	-5.0 (0.7)	-1.8 (0.4)
NB32 – placebo, Difference of LS mean	-4.8 (-5.6 to -4.0)		-4.6 (-5.2 to -3.9)		-4.2 (-5.6 to -2.9)		-3.3 (-4.3 to -2.2)	
No of patients with ≥ 5% decrease in weight, n (%)	226 (48.0)	84 (16.4)	459 (55.6)	80 (17.5)	320 (66.4)	82 (42.5)	118 (44.5)	30 (18.9)
Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	4.9 (3.6 to 6.6)		6.6 (5.0 to 8.8)		2.9 (2.0 to 4.1)		3.4 (2.2 to 5.5)	
No of patients with ≥ 10% decrease in weight, n (%)	116 (24.6)	38 (7.4)	225 (27.3)	32 (7.0)	200 (41.5)	39 (20.2)	49 (18.5)	9 (5.7)
Patients with ≥ 10% decrease in weight, NB32 vs placebo, OR (95% CI)	4.2 (2.8 to 6.2)		5.4 (3.6 to 8.0)		2.9 (2.0 to 4.4)		3.8 (1.8 to 7.9)	

Source: Section 4.7 of the CS[1]
Footnote: *Week 28 results
BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus

Table 4.9: Comparison of mITT and ITT results: percentage weight loss

Type of analysis	Outcome	COR-I[25]		COR-II[27]*		COR-BMOD[26]		COR-DM[28]	
		NB32	Pbo	NB32	Pbo	NB32	Pbo	NB32	Pbo
mITT	Percent change from baseline at end of study, LS mean (SE)	-6.1 (0.3)	-1.3 (0.3)	-6.5 (0.2)	-1.9 (0.3)	-9.3 (0.4)	-5.1 (0.6)	-5.0 (0.7)	-1.8 (0.4)
	NB32 – placebo, Difference of LS mean	-4.8 (-5.6 to -4.0)		-4.6 (-5.2 to -3.9)		-4.2 (-5.6 to -2.9)		-3.3 (-4.3 to -2.2)	
ITT using weight regain imputation method (ITT Imp)	Percent change from baseline at end of study, LS mean (SE)	-4.6 (0.3)	-1.2 (0.3)	-5.2 (0.2)	-1.9 (0.3)	NR	NR	-3.5 (0.3)	-1.7 (0.4)
	NB32 – placebo, Difference of LS mean	-3.4 (-4.1 to -2.7)		-3.4 (-3.9 to -2.8)		NR		-1.9 (-2.8 to -0.9)	
ITT using baseline-carried forward analysis (ITT BOCF)	Percent change from baseline at end of study, LS mean (SE)	-4.0 (0.3)	-0.9 (0.3)	-4.8 (0.2)	-1.5 (0.3)	-5.9 (0.4)	- 4.0 (0.6)	-3.1 (0.3)	-1.3 (0.4)
	NB32 – placebo, Difference of LS mean	-3.1 (-3.8 to -2.4)		-3.3 (-3.9 to -2.7)		-1.9 (-3.2 to -0.6)		-1.7 (-2.7 to -0.8)	

Source: Section 4.7 of the CS[1] and CS appendices[38]

Footnote *Week 28 results

BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus, Imp = weight regain imputation, BOCF = baseline observation carried forward

Table 4.10: Comparison of mITT and ITT results: patients with ≥ 5% decrease in weight

Type of analysis	Outcome	COR-I[25]		COR-II[27]*		COR-BMOD[26]		COR-DM[28]	
		NB32	Pbo	NB32	Pbo	NB32	Pbo	NB32	Pbo
mITT	No of patients with ≥ 5% decrease in weight, n (%)	226 (48.0)	84 (16.4)	459 (55.6)	80 (17.5)	320 (66.4)	82 (42.5)	118 (44.5)	30 (18.9)
	Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	4.9 (3.6 to 6.6)		6.6 (5.0 to 8.8)		2.9 (2.0 to 4.1)		3.4 (2.2 to 5.5)	
ITT Imp	No of patients with ≥ 5% decrease in weight, n (%)	203 (34.8)	78 (13.4)	446 (44.6)	79 (16.0)	NR	NR	104 (31.0)	27 (15.9)
	Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	3.6 (2.7 to 4.9)		4.7 (3.5 to 6.2)		NR			
ITT BOCF	No of patients with ≥ 5% decrease in weight, n (%)	180 (30.9)	67 (11.5)	421 (42.1)	69 (13.9)	NR	NR	94 (28.1)	24 (14.1)
	Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	3.6 (2.6 to 4.9)		4.9 (3.7 to 6.6)		NR		2.4 (1.4 to 3.9)	

Source: Section 4.7 of the CS[1] and CS appendices[38]

Footnote: *Week 28 results

BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus, Imp = weight regain imputation, BOCF = baseline observation carried forward

Subgroup analyses

Subgroup analyses were conducted for the co-primary efficacy variable (percentage of change in total body weight and proportion of patients with ≥5% decrease in total body weight at week 56 or week 28). Subgroups included study centre, sex, race, age, age group, BMI category, and tobacco use inter alia. The company did not provide results data on subgroups in the main submission document but stated that for all four trials results were 'generally consistent' or 'consistent' with the main findings. The percentages of overweight patients (BMI < 30 kg/m^2) in the trials are too small to present meaningful subgroup analyses. As the ERG had identified that the majority of participants in the trials were women we present the subgroup results for males and females separately.

Table 4.11: Results in male subgroups (mITT data)

	COR-I[25, 34]		COR-II[27, 36]		COR-BMOD[26, 35]		COR-DM[28, 37]	
	NB32	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
N	76	80	NR		56	17	121	75
Baseline weight, mean kg (SD)	115.36 (18.58)	112.16 (14.35)			118.11 (16.09)	122.12 (18.40)	116.83 (17.72)	111.89 (16.33)
End of study weight, mean kg (SD)	109.16 (18.10)	110.09 (15.33)			107.96 (18.99)	116.94 (20.35)	111.27 (18.58)	110.09 (15.95)
Percent change from baseline at end of study, LS mean (SE)	-5.20 (0.68)	-1.83 (5.92)			-8.75 (0.93)	-4.75 (1.70)	-4.79 (0.47)	-1.51 (0.60)
NB32 – placebo, Difference of LS mean (SE)	-3.34 (0.94)				-4.00 (1.94)		-3.28 (0.77)	
No of patients with ≥ 5% decrease in weight, n (%)	29 (38.16)	16 (20.00)			39 (69.64)	7 (41.18)	51 (42.15)	10 (13.33)
Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	2.36 (1.14, 4.86)				3.12 (0.99, 9.80)		4.69 (2.19, 10.05)	

Source: Trial CSRs

BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus

Table 4.12: Results in female subgroups (mITT data)

	COR-I[25, 34]		COR-II[27, 36]		COR-BMOD[26, 35]		COR-DM[28, 37]	
	NB32	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
N	395	431	NR		426	176	144	84
Baseline weight, mean kg (SD)	97.24 (14.02)	96.90 (13.00)			98.40 (13.81)	99.55 (13.19)	97.55 (15.49)	98.82 (15.47)
End of study weight, mean kg (SD)	91.29 (15.72)	95.80 (14.11)			88.79 (15.58)	94.40 (15.39)	92.31 (16.09)	96.71 (16.11)
Percent change from baseline at end of study, LS mean (SE)	-6.23 (0.34)	-1.15 (0.32)			-9.83 (0.41)	-5.66 (0.64)	-5.41 (0.46)	-2.15 (0.60)
NB32 – placebo, Difference of LS mean (SE)	-5.08 (0.46)				-4.17 (0.76)		-3.25 (0.76)	
No of patients with ≥ 5% decrease in weight, n (%)	197 (49.87)	68 (15.78)			281 (65.96)	75 (42.61)	67 (46.53)	20 (23.81)
Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	5.37 (3.87, 7.44)				2.59 (1.81, 3.71)		2.77 (1.52, 5.05)	

Source: Trial CSRs

BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus

It can be seen from Table s 4.11 and 4.12 that both men and women taking NB32 have statistically significantly better results than those taking placebo.

ERG comment:

- Based on the mITT data presented by the company NB32 results in greater weight loss and in a higher number reporting 5% or more weight loss. However the ERG draws the attention of the committee to the superior results of the BMOD trial. NB32 together with a more intensive behaviour modification programme resulted in 66.4% of patients losing 5% or more weight compared to 44 to 55% in the other three trials without such an intensive intervention. Moreover, in the BMOD trial the placebo and behaviour modification arm achieved results approaching the medication arms in the other trials.
- Both men and women appear to benefit from the intervention when comparing subgroups.
- Using the true ITT data, NB32 also results in a greater mean percentage of weight loss compared to placebo groups. The proportion of patients losing 5% or more weight is also higher in treatment groups. However the results, as expected, are lower than the mITT data. It is these data that should be used to determine effectiveness and to ascertain the clinical importance of the results.

Table 4.13 shows the reasons for treatment discontinuation across the trials.

Table 4.13: Reasons for treatment discontinuation

	COR-I[25]		COR-II[27]		COR-BMOD[26]		COR-DM[28]	
	NB32	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
Number randomised	583	581	1001	495	591	202	335	170
Number discontinued (%)	287 (49%)	291 (50%)	462 (46%)	226 (46%)	249 (42%)	84 (42%)	160 (48%)	70 (41%)
Reasons for discontinuation								
Adverse event	112	56	241	68	150	25	98	26
Lost to follow up	65	66	77	48	22	17	22	15
Withdrew consent	60	90	75	56	43	24	21	15
Enrolled but did not meet criteria	0	1	0	0	0	0	2	0
Study drug not dispensed	0	3	0	0	3	0	0	1
Participant judged weight loss insufficient	12	40	19	33	3	6	5	6
Drug non-compliance	17	15	31	13	13	5	8	3
Protocol non-compliance	9	7			4	0	3	4
Death	1	0	0	0	0	0	0	0
Other	11	13	19	8	11	7	1	0

Source: Figures 3 to 6 of the CS and accompanying text[1]
BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus

In COR-I, 870 (50%) of patients completed 56 weeks of treatment. 287 of 583 (49%) in the NB32 group discontinued whilst 291 of 581 (50%) in the placebo group discontinued. The company stated that '*Rates of discontinuation were similar across treatment groups*.'[1] However they noted (as can be seen in Table 4.13) more patients in the NB32 group discontinued due to adverse events than in the placebo group ($p < 0.0001$). More patients in the placebo group discontinued due to insufficient weight loss (6.9% vs. 2.1%, $p < 0.0001$) and withdrawal of consent (15.5% vs. 10.3%, $p = 0.0126$). The company stated that rates of discontinuation were higher during the first 16 weeks of the study in both treatment and placebo groups. A similar pattern was observed in COR-II where 46% of patients in each treatment group discontinued during 56 weeks of treatment. More NB-32 treated patients discontinued due to an adverse event (24.1% vs. 13.7%, $p < 0.001$) whilst more placebo group patients discontinued due to insufficient weight loss (6.7% vs. 1.9%, $p < 0.001$) and withdrawal of consent (11.3% vs. 7.5%, $p < 0.05$). In COR-BMOD 41.6% of the placebo group and 42.1% of the NB32 group discontinued treatment. Again a greater percentage of those in the NB32 group discontinued due to an adverse event (25.4% vs. 12.4%, $p < 0.001$). A greater number of placebo patients discontinued due to withdrawal of consent (11.9% vs. 7.3%, $p = 0.042$), loss to follow up (8.4% vs. 3.7%, $p = 0.008$) or self-perceived insufficient weight loss (3.0% vs. 0.5%, $p = 0.004$). In COR-DM 47.8% of the NB32 treatment group discontinued treatment compared with 41.2% in the placebo group. Again, a greater percentage of patients who received NB32 compared with placebo discontinued due to an adverse event (29.3% vs. 15.3%) A greater percentage in the placebo group withdrew as they were lost to follow up (8.8% vs.6.6%), withdrew consent (8.8% vs. 6.3%) or had self-perceived insufficient weight loss (3.5% vs. 1.5%).

ERG comment:

- Across the four main trials treatment discontinuation rates ranged from approximately 41 to 50%. This suggests that in practice up to half of patients may complete a year's treatment with NB32. Rates of discontinuation were found to be similar between NB32 and placebo in all trials.
- Reasons for discontinuation varied between treatment groups. In all four trials a higher number of patients in NB32 groups discontinued due to an adverse event. As more NB32 patients in each trial discontinued due to an adverse event this indicates that the mITT population is likely to be biased as these patients would be more likely to be missing a post-baseline weight assessment and be excluded from the mITT analysis. Adverse events will be discussed further in Section 4.2.6 of this report. Although the placebo groups in all trials had more participants discontinuing due to insufficient weight loss, percentages of patients citing this reason were relatively low (approximately 7% in COR-I and COR-II, 3% in COR-BMOD and 3.5% in COR-DM).

4.2.5 Direct evidence: Meta-analysis results

The company reported the results of meta-analyses for the NB32 trials in Chapter 4.9.3 of the CS (CS, Figures 16 and 17, pages 115-117) and results of sensitivity analyses are presented in Appendix 8 of the CS.

The four trials COR-I, COR-II, COR-BMOD and COR-DM were pooled in random effects meta-analyses for the two weight loss outcomes. However the moderate to high levels of statistical heterogeneity observed ($I^2 = 66.6\%$ for $\geq 5\%$ reduction in weight and 70.1% for percentage weight change) indicate variation in the results between the trials. Sensitivity analyses pooled the T2DM trial (COR-DM) and non-T2DM trials (COR-BMOD, COR-I and COR-II) separately which reduced this

heterogeneity to $I^2 = 0\%$ for the percentage weight change analysis but not for the ≥ 5% reduction in weight analysis where the heterogeneity remained high. Further analyses then excluded the COR-BMOD trial which removed the observed heterogeneity from the ≥ 5% reduction in weight analysis.

ERG comment: The trials were conducted in two different populations (with and without T2DM) and one trial used a more intensive behaviour modification as the background therapy (COR-BMOD) compared to the other three trials. As these differences in populations and interventions appear to be linked to the statistical heterogeneity between the results they should be pooled separately. The COR-I and COR-II trials are clinically similar and could in theory be pooled for the no T2DM analysis. However, as COR-II assessed results at 28 weeks and patients were re-randomised after this it should not be pooled with COR-I. COR-DM is the only trial available for the T2DM analysis and COR-BMOD should be considered separately due to the more intensive behaviour management therapy. Therefore, the ERG believes none of the NB32 trials should be pooled.

4.2.6 Direct evidence: Safety results

Safety results were based on the four main trials: COR-I, COR-II, COR-BMOD and COR-DM. According to the CS treatment-emergent adverse events were defined as "*events that first occurred or worsened during double-blind treatment (i.e. a new event or an exacerbation of a pre-existing condition) with an onset date after study drug administration and within 7 days of the last confirmed dose date. AEs with an onset date before the first dose of study drug were recorded under medical history*."[1] Events were categorised across the trials as mild, moderate or severe and relationship to the study drug was investigated.

The company further stated that "*Safety data are presented for the safety analysis set, defined as all randomised patients who were administered at least one tablet of study treatment and had at least one investigator contact/assessment at any time after the start of study treatment, regardless of whether they discontinued the study. Patients were grouped in the safety analysis set according to which study treatment was administered on the first day of treatment following randomisation.*"[1]

An overview of adverse events is presented in Table 4.14.

Table 4.14: Overview of adverse events

	COR-I[25]		**COR-II[27]**		**COR-BMOD[26]**		**COR-DM[28]**	
	NB32	**Pbo**	**NB32 /48**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
Safety analysis set								
All TEAEs, n (%)	476 (83.1)	390 (68.5)	852 (85.9)	370 (75.2)	547 (93.7)	176 (88.0)	301 (90.4)	144 (85.2)
Drug-related TEAEs, n (%)	336 (58.6)	167 (29.3)	630 (63.5)	189 (38.4)	447 (76.5)	108 (54.0)	238 (71.5)	57 (33.7)
Severe TEAEs, n (%)	51 (8.9)	34 (6.0)	110 (11.1)	33 (6.7)	98 (16.8)	15 (7.5)	61 (18.3)	19 (11.2)
TESAEs, n (%)	9 (1.6)a	8 (1.4)a	21 (2.1)	7 (1.4)	22 (3.8)	1 (0.5)	13 (3.9)	8 (4.7)
Discontinued due to AEs, n (%)	112 (19.5)	56 (9.8)	241 (24.3)	68 (13.8)	150 (25.7)	25 (12.5)	98 (29.4)	26 (15.4)
Deaths	1	0	0	0	0	0	0	0

	COR-I[25]	COR-II[27]	COR-BMOD[26]	COR-DM[28]
Source: Tables 42, 44, 46 and 48 of the CS[1] and Trial CSRs[34-37] Footnote: a) none found to be related to the drug AE = adverse event; BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus; TEAE = treatment emergent adverse event; TESAE = treatment emergent serious adverse event				

Adverse events occurred in 83.1% to 93.7% of treatment groups and 68.5% to 88.0% of placebo groups. Approximately 58% to 76% of these were attributed to the drug in NB32 groups across the trials. Serious adverse events occurred at similar rates in treatment and placebo groups across the trials. However a larger number of patients discontinued due to adverse events across the trials (19.5% to 29.4% for treatment groups vs. 9.8% to 15.4% in placebo groups).

The main specific adverse events are listed in Table 4.15.

Table 4.15: Specific adverse events (≥ 5% in at least one treatment arm of an included trial)

	COR-I[25]		COR-II[27]		COR-BMOD[26]		COR-DM[28]	
	NB32 (n = 573)	**Pbo (n = 569)**	**NB32/ 48 (n = 992)**	**Pbo (n = 492)**	**NB32 (n = 584)**	**Pbo (n =200)**	**NB32 (n = 333)**	**Pbo (n = 169)**
Adverse Event, n (%)								
Gastrointestinal disorders	**292 (51.0)**	**136 (23.9)**	**532 (53.6)**	**131 (26.6)**	**380 (65.1)**	**78 (39.0)**	**215 (64.6)**	**53 (31.4)**
Nausea	171 (29.8)	30 (5.3)	290 (29.2)	34 (6.9)	199 (34.1)	21 (10.5)	141 (42.3)	12 (7.1)
Vomiting	56 (9.8)	14 (2.5)	84 (8.5)	10 (2.0)	64 (11.0)	13 (6.5)	61 (18.3)	6 (3.6)
Constipation	90 (15.7)	32 (5.6)	189 (19.1)	35 (7.1)	141 (24.1)	28 (14.0)	59 (17.7)	12 (17.1)
Dry mouth	43 (7.5)	11 (1.9)	90 (9.1)	13 (2.6)	47 (8.0)	6 (3.0)	21 (6.3)	5 (3.0)
Diarrhoea	26 (4.5)	28 (4.9)	55 (5.5)	18 (3.7)	43 (7.4)	15 (7.5)	52 (15.6)	16 (9.5)
Abdominal pain upper	NR	NR	NR	NR	32 (5.5)	3 (1.5)	17 (5.1)	3 (1.8)
Infections and infestations	**203 (35.4)**	**200 (35.1)**	**359 (36.2)**	**205 (41.7)**	**188 (32.2)**	**63 (31.5)**	**121 (36.3)**	**77 (45.6)**
Upper respiratory tract infection	57 (9.9)	64 (11.2)	86 (8.7)	55 (11.2)	38 (6.5)	18 (9.0)	26 (7.8)	16 (9.5)
Sinusitis	30 (5.2)	34 (6.0)	51 (5.1)	35 (7.1)	16 (2.7)	6 (3.0)	16 (4.8)	14 (8.3)
Nasopharyngitis	29 (5.1)	31 (5.4)	82 (8.3)	40 (8.1)	36 (6.2)	15 (7.5)	28 (8.4)	23 (13.6)
Musculoskeletal and connective tissue disorders	**72 (12.6)**	**90 (15.8)**	**159 (16.0)**	**96 (19.5)**	**104 (17.8)**	**46 (23.0)**	**58 (17.4)**	**40 (23.7)**
Nervous system disorders	**167 (29.1)**	**95 (16.9)**	**326 (32.9)**	**81 (16.5)**	**263 (45.0)**	**60 (30.0)**	**129 (38.7)**	**32 (18.9)**

	COR-I[25]		COR-II[27]		COR-BMOD[26]		COR-DM[28]	
Headache	79 (13.8)	53 (9.3)	174 (17.5)	43 (8.7)	139 (23.8)	35 (17.5)	46 (13.8)	15 (8.9)
Dizziness	54 (9.4)	15 (2.6)	68 (6.9)	18 (3.7)	85 (14.6)	9 (4.5)	390 (11.7)	9 (5.3)
Tremor	12 (2.1)	1 (0.2)	35 (3.5)	3 (0.6)	34 (5.8)	2 (1.0)	22 (6.5)	4 (2.4)
Psychiatric disorders	**85 (14.8)**	**62 (10.9)**	**205 (20.7)**	**75 (15.2)**	**145 (24.8)**	**45 (22.5)**	**75 (22.5)**	**20 (11.8)**
Insomnia	43 (7.5)	29 (5.1)	97 (9.8)	33 (6.7)	51 (8.7)	12 (6.0)	37 (11.1)	9 (5.3)
Anxiety	9 (1.6)	12 (2.1)	48 (4.8)	21 (4.3)	30 (5.1)	7 (3.5)	18 (5.4)	2 (1.2)
Vascular disorders	**51 (8.9)**	**22 (3.9)**	**(7)**	**(3.3)**	**46 (7.9)**	**7 (3.5)**	**40 (12.0)**	**12 (7.1)**
Hot flush	30 (5.2)	7 (1.2)	42 (4.2)	6 (1.2)	28 (4.8)	1 (0.5)	7 (2.1)	4 (2.4)
Tinnitus	15 (2.6)	6 (1.1)	29 (2.9)	1 (0.2)	31 (5.3)	1 (0.5)	8 (2.4)	1 (0.6)
Hypertension	17 (3.5)	14 (2.5)	(1.9)	(1.6)	14 (2.4)	4 (2.0)	33 (9.9)	7 (4.1)

Source: Tables 43, 45, 47 and 49 of the CS[1] and Trial CSRs[34-37]
Footnote: Adverse event categories in bold

The main category of adverse event occurring more frequently in treatment groups across the trials was gastrointestinal disorders. Nausea, in particular, occurred frequently and more often in treatment groups. Across the trials rates of nausea ranged from 29.2% to 42.3% in treatment groups. Rates ranged from 5.3% to 10.5% in placebo groups. Vomiting, constipation and dry mouth also occurred more frequently in treatment groups although at a lower rate than that of nausea. Nervous system disorders such as headache, dizziness and tremor occurred more frequently in treatment groups.

The incidence of events of particular concern (serious cardiovascular disorders and suicidality measured on IDS) was extremely small and any differences between groups could not be ascertained in view of the small numbers in both groups.

ERG comment:

- The ERG draws to the attention of the committee the greater proportion of gastrointestinal events, particularly nausea, in NB32 groups across the trials. Although the majority of events were not serious, more participants withdrew as a result of adverse events in treatment groups. This finding is relevant to implementation of the intervention in clinical practice.
- The ERG notes that the NB-CVOT trial (described in Section 4.2.7 of the report below) was primarily designed to investigate the cardiovascular safety of NB32 in weight management. However the study was terminated earlier than originally planned (after the 50% interim analysis), after interim data were made public in a US patent (and related Orexigen security filings) and by the EMA in the Mysimba EPAR.
- A further trial on occurrence of MACE in overweight and obese patients with cardiovascular disease receiving NB32 was requested by CHMP. Based in the US, this randomised trial will

aim to enrol 8,000 patients and is estimated to last for up to six years until the targeted number of adjudicated MACE events (378) has been reached. The primary MACE composite comprises the first occurrence of CV death, nonfatal myocardial infarction (MI), and nonfatal stroke.[31]

4.2.7 Overview of the supporting RCTs

As previously stated, two trials (NB-CVOT and IGNITE) were used to provide data on long-term safety and efficacy only. NB-CVOT was presented in detail in the submission. IGNITE was presented only briefly and the company stated that '*At the time of database searches, this study was not yet published and was therefore not identified or included in the SLR*'.[1] This section will give an overview of each trial in turn.

NB-CVOT

The CS stated that "*The NB-CVOT study was a Phase IIIb, multicentre, randomised, double-blind, placebo-controlled trial to assess the occurrence of MACE in overweight or obese patients*."[1] Study details are given in the table below.

Table 4.16: Overview of NB-CVOT

Participants	Intervention (n =4,454)	Control (n =4,456)	Trial design and duration	Primary outcome
Patients aged 45 (men) or 50 (women) years or older, with a BMI 27–50kg/m^2 and a waist circumference of 88cm (women) or 102cm (men) or more. Patients had characteristics associated with an increased risk of adverse CV outcomes.	Naltrexone 32 mg per day + bupropion 360 mg per day (NB32) + customary diet and behaviour modification Patients were also encouraged to participate in an internet-based weight management program as well as having access to a personal weight loss coach and a low-fat, low-calorie meal plan.	Placebo + customary diet and behaviour modification Patients were also encouraged to participate in an internet-based weight management program as well as having access to a personal weight loss coach and a low-fat, low-calorie meal plan.	Lead-in period 4 week dose escalation period Maintenance period At 16 weeks, if patients did not lose ≥2% of their initial body weight or experienced a sustained (at ≥2 visits) increase in blood pressure (systolic or diastolic) of 10mmHg or greater they were discontinued	Time from treatment randomisation to the first confirmed occurrence of a MACE, defined as CV death, nonfatal stroke, or nonfatal myocardial infarction. Only SAEs and AEs leading to study drug discontinuation were collected. No subgroup analyses

Source: Table 11 and section 4.3 of the CS[1]
AE = adverse event; CV = cardiovascular; MACE = major adverse cardiovascular events; SAE = serious adverse event

The main differences between the NB-CVOT trial and the COR trials is that participants were all at increased risk of adverse cardiovascular outcomes. Furthermore, the trial incorporated a lead-in period. During the lead in period 1,490 patients discontinued. Of these 543 discontinuations were due to adverse events.

The trial incorporated a stopping rule at 16 weeks (unlike the COR trials). The company stated that "*A large decrease in the number of patients receiving the study drug occurred after the 16-week assessment, with 44% of placebo patients and 17.8% of NB32 patients discontinued by investigators.*

Most discontinuations were due to a failure to lose 2% of body weight, but 230 placebo patients and 154 NB32 patients discontinued treatment because of a greater than 10mmHg increase in blood pressure. A high percentage of patients who discontinued treatment remained in follow-up for MACE and contributed to the ITT analysis set."[38]

NB-CVOT was terminated early (after the 50% interim analysis), after 25% interim data were made public in a US patent (and related Orexigen security filings) and by the EMA in the Mysimba EPAR.

Patient characteristics are given in Table 4.17.

Table 4.17: NB-CVOT patient characteristics

Patient characteristics		
	NB32	**Pbo**
Age, mean years (SD)	61.1 (7.27)	60.9 (7.38)
Age range (min, max)	45, 86	45, 85
Sex, female, n (%)	2437 (54.7)	2419 (54.4)
Ethnicity, n (%)		
White	3738 (83.9)	3698 (83.1)
Black	656 (14.7)	648 (14.6)
Asian	19 (0.4)	27 (0.6)
Other	41 (0.8)	75 (1.6)
BMI, mean kg/m^2 (SD)	37.2 (5.26)	37.4 (5.44)
Obesity class, n (%)		
BMI < 30 kg/m^2	299 (6.7)	311 (7.0)
BMI ≥30 and <35 kg/m^2	1393 (31.3)	1408 (31.6)
BMI ≥35 and <40 kg/m^2	1476 (33.2)	1383 (31.1)
BMI ≥40 kg/m^2	1284 (28.8)	1348 (30.3)
Weight, mean kg (SD)	105.6 (19.09)	106.3 (19.18)
Smoker, n (%)	405 (9.1)	416 (9.3)
Hypertension, n (%)	4162 (93.4)	4117 (92.5)
Dyslipidaemia, n (%)	4100 (92.0)	4070 (91.5)
Type 2 diabetes	3784 (84.9)	3803 (85.5)
Alcohol use, n (%)	NR	NR
History of depression	1031 (23.1)	995 (22.4)
History of anxiety	NR	NR
Source: CS[1] and NB-CVOT CSR[31]		

The mean age of randomised patients was 61 years compared to approximately 44 years in the COR trials (54 in COR-DM). Just over half are female (similar to COR-DM) whereas the majority are female in COR-I, COR-II and COR-BMOD. Ethnicity was similar to the COR trials with the majority of participants being white.

In terms of comorbidities, T2DM was present in 85.2% of patients (0 in the COR trials except for COR-DM (100%)). 32.1% had cardiovascular disorders but the company described cardiovascular risk factors as "*well-controlled*".[1] Nearly all patients in NB-CVOT had hypertension or dyslipidaemia, Concomitant

medications included statins in 79.2% of patients, anti-hypertensive medications in 92.0%, and glucose lowering agents in 75.1%, and anti-depressant medication in 23.1% of patients. The company stated that there was a higher instance of depression in NB-CVOT than in the COR trials.[1]

The CS stated that "*The 50% interim analysis was completed on 3 March 2015, based on 192 adjudicated major adverse cardiovascular events (MACE) (from a database lock on 3 February 2015). Additional outcomes accumulated after the February 2015 database lock are included in a sensitivity analysis, which reports results after 64% of planned events."* Results are presented in the Table below for the 50% analysis.

Table 4.18: NB-CVOT effectiveness outcomes

	NB32 (n=4,455)	Placebo (n=4,450)	HR (99.7% CI)
Primary outcome, n (%)			
MACE	90 (2.0)	102 (2.3)	0.9 (0.6–1.3)
CV death	17 (0.4)	34 (0.8)	0.5 (0.2–1.2)
Nonfatal stroke	21 (0.5)	19 (0.4)	1.1 (0.4–2.8)
Nonfatal myocardial infarction	54 (1.2)	54 (1.2)	1.0 (0.6–1.8)
Secondary outcomes, n (%)			
MACE + hospitalisation for unstable angina	133 (3.0)	142 (3.2)	0.9 (0.7–1.3)
Fatal or nonfatal stroke	22 (0.5)	21 (0.5)	1.0 (0.4–2.6)
Fatal or nonfatal myocardial infarction	55 (1.2)	57 (1.3)	1.0 (0.6–1.7)
Other outcomes, n (%)			
All-cause mortality	43 (1.0)	51 (1.1)	0.8 (0.5–1.5)
Hospitalisation for unstable angina	47 (1.1)	43 (1.0)	1.1 (0.6–2.0)
Coronary revascularisation	132 (3.0)	145 (3.3)	0.9 (0.6–1.3)
All-cause mortality + nonfatal myocardial infarction + nonfatal stroke	114 (2.6)	119 (2.7)	1.0 (0.7–1.4)
MACE + hospitalisation for unstable angina + coronary revascularisation	188 (4.2)	205 (4.6)	0.9 (0.7–1.2)

Source: CS[1]
CV = cardiovascular; MACE = major adverse cardiovascular events

For the 50% interim analysis, time to first MACE, occurred in 192 patients; 102 (2.3%) in the placebo group and 90 (2.0%) in the NB32 group (HR: 0.88; 99.7% CI: 0.57, 1.34). The components of the primary composite outcome included CV death (0.8% of placebo patients and 0.4% of NB32 patients [HR: 0.50; 99.7% CI: 0.21, 1.19]), nonfatal stroke (0.4% of placebo patients and 0.5% of NB32 patients [HR: 1.10; 99.7% CI: 0.44, 2.78]) and nonfatal myocardial infarction (1.2% in both placebo and NB32 patients [HR: 1.00; 99.7% CI: 0.57, 1.75]). The company stated that "*In general, final end-of-study analyses support these data*." (CS, page 109).

The company further stated that "*At trial completion, body weight decreased by a mean of 3.9kg (95% CI: -4.1, -3.7kg) in the NB32 group compared to a mean decrease of 1.2kg (95% CI: -1.3, -1.0kg) in the placebo group, corresponding to reductions of 3.6% and 1.1%, respectively (p<0.001)*." (CS, page 109).

Adverse events identified in NB-CVOT are detailed in Table 4.19.

Table 4.19: NB-CVOT adverse events

	NB32 (n=4455)	Placebo (n=4450)
Drug-related TEAEs, n (%)	982 (22.0)	174 (3.9)
Severe TEAEs, n (%)	217 (4.9)	108 (2.4)
TESAEs, n (%)	463 (10.4)	386 (8.7)
DC due to AEs, n (%)	1292 (29.0)	400 (9.0)
Deaths, n (%)	65	72
Source: CS[1] AE = adverse event; DC = discontinuation; TEAE = treatment emergent adverse event; TESAE = treatment emergent serious adverse event.		

It can be seen from Table 4.19 that more patients in the NB32 group experienced events that were considered by the investigator to be study drug–related (22.0% vs. 3.9% with placebo). In both groups, most TEAEs leading to discontinuation were considered mild or moderate in intensity. TESAEs, (defined as any AE occurring at any dose of study drug that resulted in death, life-threatening adverse drug experience, inpatient hospitalisation or prolongation of existing hospitalisation, persistent of significant disability or incapacity, important medical events or congenital anomaly or birth defect) were reported for 849 patients (9.5%) overall, 10.4% in the NB32 group and 8.7% in the placebo group. A total of 137 deaths occurred during the study, 65 patients in the NB32 group and 72 in the placebo group, although no deaths in this study were related to the study drug.

Discontinuations due to adverse events most commonly included gastrointestinal AEs, which occurred in 14.2% of NB32 patients and 1.9% of placebo-treated patients and central nervous system symptoms, which occurred in 5.1% of NB32 patients and 1.2% of placebo patients. Psychiatric symptoms resulted in study drug discontinuation in 3.1% of NB32 patients and 0.9% of placebo patients ($p<0.001$).

IGNITE

IGNITE was described in the CS as a "*Phase IIIb, randomised, open label, controlled study in which patients received NB32 plus comprehensive lifestyle intervention (CLI) or usual care (standard diet and exercise advice) for 26 weeks.*"[1] Patients in the NB32 + CLI group not achieving 5% weight loss at week 16 were discontinued. After week 26 patients in the usual care arm began NB32 + CLI and were assessed up to week 78. The primary endpoint was percentage change in weight from baseline to week 26 in the per protocol (PP) population. Other endpoints included percentage of patients achieving ≥5%, ≥10% and ≥15% weight loss, percent change in weight at week 78 and AEs necessitating study discontinuation. No subgroup analyses were conducted.

A total of 242 patients were randomised; 153 to NB32 + CLI and 89 patients to usual care for a total of 26 weeks. It was noted in the CS that although the trial was of 78 weeks' duration, patient numbers beyond 52 weeks were low; 61 NB32 patients were followed from week 52 onwards.

The CS stated that *"Patients assigned to treatment with NB32 plus standard management lost significantly more weight than patients treated with usual care at Week 26 (-9.4% vs -0.94% respectively; p<0.0001). For patients who remained on treatment, the initial weight loss observed at 26 weeks was sustained throughout Week 78, further supporting the maintained effectiveness of NB32 treatment."*

The CS further stated that *"In the IGNITE study, the safety profile shown was consistent with that seen in the previous, pivotal trials; most patients tolerated NB32 well, and those who developed AEs did so early in the treatment protocol. The most common AE leading to NB discontinuation was nausea (7.0% of all subjects), which is consistent with the rate in the Phase III trials (6.3%)."*

No further information was provided on IGNITE.

ERG comment:

- The NB-CVOT study has the potential to provide information on performance of NB32 in an older population with cardiovascular disease when compared to the COR trials. Most of the patients in NB-CVOT are diabetic, and many are depressed. However a number of problems were identified with NB-CVOT. These include the use of a lead-in period where large numbers of patients discontinued primarily due to adverse events. This implies that those continuing to the treatment period who were re-randomised were better able to tolerate the drug. The adverse event profile will be an overestimate of the tolerability of the drug. In addition only SAEs and AEs leading to study drug discontinuation were collected. Even so, an elevated number of gastrointestinal events were noted in the NB32 group. A further limitation is that the trial was terminated early (after the 50% interim analysis), after 25% interim data were made public in a US patent (and related Orexigen security filings) and by the EMA in the Mysimba EPAR. The trial was not able to provide a definitive answer to the cardiovascular risk of NB32 and a further trial has been instigated. The reliability of the final data on weight loss is also questionable.
- The IGNITE trial was described only briefly in the submission. The main limitation of this trial was that intervention and control groups were not directly comparable. This trial is not able to assess the unique effect of NB32 but only the combined effect of NB32 and a comprehensive lifestyle intervention. Furthermore the trial was only randomised for 26 weeks rather than 56 for most of the COR trials.

4.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

No trials were identified that compared NB32 directly with orlistat. Therefore, the company performed indirect comparisons to compare NB32 with orlistat using placebo as the common comparator.

As described in Section 4.1.2 of this report, the search was not aimed at finding studies of behavioural or lifestyle interventions. In addition to the inclusion criteria described in Section 4.1.2 of this report, the following studies/treatment arms were excluded:

- Treatment arm is not of interest
- Treatment group is not administered at recommended dosage
- Trial reduces to single treatment arm once other arms are pooled or excluded
- Trial reports no relevant outcome data
- Trial excludes patients during a lead-in period due to weight loss criteria or treatment compliance
- Trial has a wait list control group as a comparator arm in which patients receive no pharmaceutical treatment or standard management

For the analyses performed in the CS, NB and orlistat were evaluated at their recommended doses detailed in the summary of product characteristics for each treatment[39, 40]:

- NB – naltrexone 32mg/day prolonged release plus bupropion 360mg/day prolonged release (NB32)

- Orlistat – 120mg three times a day (TID)

ITC were performed to compare NB32 and orlistat for the following outcomes:

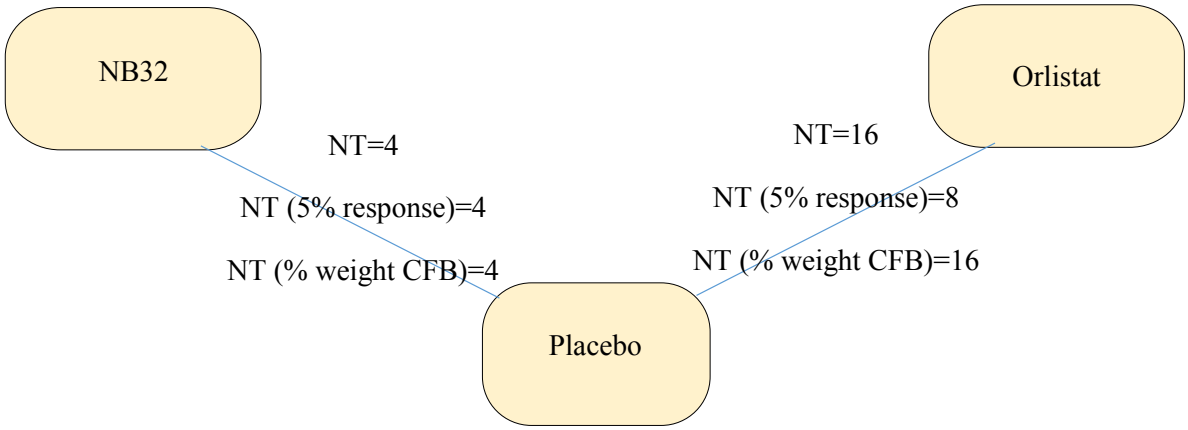
- Mean percentage weight change from baseline at one year (the one year time point ranged from 52 to 57 weeks [continuous outcome])
- At least 5% reduction in weight at one year from baseline (the one year time point ranged from 52 to 57 weeks [dichotomous outcome])

According to the CS, 36 RCTs were included (CS, page 46) in the systematic literature review, five studies investigated treatment with NB32 (COR-I, COR-II, COR-BMOD, COR-DM and NB-CVOT), while the remaining 31 studies investigated treatment with orlistat.

As explained in Section 4.2 of this report, four NB32 studies were used in the indirect comparisons (COR-I, COR-II, COR-BMOD and COR-DM); NB-CVOT was excluded from the analyses due to the trial design, objective, and patient population, being different from the other NB32 studies and patients were excluded during the lead in period. In addition, 16 out of 31 orlistat studies were used in the indirect comparisons. Reasons for exclusion of the 15 orlistat trials not used in the analyses are explained in Appendix 9 of the CS.

Twenty trials were included in the indirect treatment comparison (ITC). The list of trials, along with treatments and available outcome data that were included in the analyses are presented in Table 4.20. The maximum evidence base for each outcome, following the additional exclusion is given in the network of evidence presented in Figure 4.1.

Figure 4.1: Network of evidence



Source: CS, Figure 18, page 124[1]
Notes**:** 5% response defined as ≥5% reduction in weight from baseline at 1 year.
CFB = change from baseline; NB32 = naltrexone 32mg plus bupropion; NT = number of trials.

Table 4.20: Evidence base: trials, treatments and outcomes included in the ITC analyses

Trial (NT=20)	Trial duration	Treatment		Analysis population					Outcome	
		Arm 1	Arm 2	Trials without a lead-in period	Trial excludes T2DM patients[a]	T2DM is part of trial inclusion criteria	Trials without a high proportion of comorbidities[b]	StMan without intensive BMOD[c]	≥5% reduction in weight	Mean % weight CFB
Apovian 2013[27] (COR-II; NCT00567255)[d]	56 weeks	PBO	NB32	✓	✓	-	✓	✓	✓	✓
Greenway 2010[25] (COR-I; NCT00532779)	56 weeks	PBO	NB32	✓	✓	-	✓	✓	✓	✓
Hollander 2013[28] (COR-DM; NCT00474630)	56 weeks	PBO	NB32	✓	-	✓	-	✓	✓	✓
Wadden 2011[26] (COR-BMOD; NCT00456521)	56 weeks	PBO	NB32	✓	✓	-	✓	-	✓	✓
Astrup 2012[41] (NN8022-1807 study group; NCT00422058 [extension study: NCT00480909)	54 weeks (2-week lead-in period and 52-week treatment phase [weeks 20–52 were part of an extension study])	PBO	ORL 120mg TID	-	✓	-	✓	✓	✓	✓
Bakris 2002[42]	52 weeks	PBO	ORL 120mg TID	✓	-	-	-	✓	✓	✓
Berne 2005[43] (OST2D study group)	54 weeks (2-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	✓	-	✓	✓	✓
Broom 2002[44] (UKM study group)	54 weeks (2-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	✓	✓	✓	✓
Derosa 2003[45]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO[e]	ORL 120mg TID[f]	-	-	-	-	✓	-	✓
Derosa 2010[46]	52 weeks	PBO	ORL 120mg TID	✓	-	✓	-	✓	-	✓

Trial (NT=20)	Trial duration	Treatment		Analysis population					Outcome	
		Arm 1	Arm 2	Trials without a lead-in period	Trial excludes T2DM patients[a]	T2DM is part of trial inclusion criteria	Trials without a high proportion of comorbidities[b]	StMan without intensive BMOD[c]	≥5% reduction in weight	Mean % weight CFB
Gotfredsen 2001[47] (EM Study-I)	52 weeks (4-week lead-in period and 48-week treatment period)	PBO	ORL 120mg TID	-	-	-	✓	✓	-	✓
Karhunen 2000[48] (EM Study-II)	108 weeks (4-week lead-in period and two 52-wk treatment periods)	PBO	ORL 120mg TID	-	✓	-	✓	✓	-	✓
Kelley 2002[49]	54 weeks (2-week screening and 52-week treatment phase)	PBO	ORL 120mg TID	✓	-	✓	-	✓	✓	✓
Lindgarde 2000[50]	54 weeks (2-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	-	✓	✓	✓
Lucas 2003[51]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	-	✓	-	✓
Mathus-Vliegen 2006[52]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO	ORL 120mg TID	-	✓	-	✓	✓	-	✓
Miles 2002[53]	54 weeks (2-wkk screening period and 52-week treatment phase)	PBO	ORL 120mg TID	✓	-	✓	-	✓	✓	✓
Reaven 2001[54]	56 weeks (4-week lead-in period and 52-week treatment period)	PBO[g]	ORL 120mg TID[g]	-	✓	-	✓	✓	-	✓
Swinburn 2005[55]	56 weeks (4-week lead-in period plus 52-week treatment period)	PBO	ORL 120mg TID	-	-	-	✓	✓	-	✓
Torgerson 2004[56] (XENDOS)	208 weeks	PBO	ORL 120mg TID	✓	✓	-	✓	-	✓	✓
Total NB32 trials				4	3	1	3	3	4	4

Trial (NT=20)	Trial duration	Treatment		Analysis population					Outcome	
		Arm 1	Arm 2	Trials without a lead-in period	Trial excludes T2DM patients[a]	T2DM is part of trial inclusion criteria	Trials without a high proportion of comorbidities[b]	StMan without intensive BMOD[c]	≥5% reduction in weight	Mean % weight CFB
Total ORL trials				5	5	4	8	15	8	16
Total trials				9	8	5	11	18	12	20

Source, CS, Table 29, pages 120-123[1]

Notes: [a]) As per the trial exclusion criteria; [b]) High proportion of comorbidities were defined as in Section 4.10.3 of the CS, [c]) Intensive BMOD defined as in Section4.10.4 of the CS; [d]) Non-responders in the Apovian 2013 trial were re-randomised to either NB32 or NB48. Non-responders who received NB48 after 32 weeks were not included in the analysis, and patients who received NB32 were double weighted in the analysis; [e]) PBO and PBO+FV have been pooled together; [f]) ORL 120mg TID and ORL120mg TID+FV have been pooled together; [g]) Trial presents arm data split by whether patients had syndrome X, and patients with/without syndrome X were pooled for each treatment.

BMOD = behaviour modification; CFB = change from baseline; COR = Contrave® obesity research; DM = diabetes mellitus; EM = European multicentre; FV = fluvastatin; NB = naltrexone plus bupropion; NB16 = naltrexone 16mg plus bupropion; NB32 = naltrexone 32mg plus bupropion; NB48 = naltrexone 48mg plus bupropion; NT = number of trials; ORL = orlistat 120mg TID; PBO = placebo; SM = Swedish Multimorbidity; StMan = Standard Management; T2DM = Type 2 diabetes mellitus; TID = three times a day; UKM = UK Mulitmorbidity; XENDOS = Xenical in the prevention of diabetes in obese subjects.

In total 24 analyses were performed (see CS Table 33, page 131), six base-case analyses for patients with T2DM, patients without T2DM and all patients (for each population a Bayesian and a Frequentist pairwise MA was performed). In addition four sensitivity analyses were performed for each population using Bayesian and Frequentist methods. Therefore, a total of 30 analyses would have been possible, but six analyses could not be performed due to insufficient data or because the evidence base was the same as in a previous analysis.

The four sensitivity analyses were: SA1-excluding trials where ≥75% of patients had at least one comorbidity; SA2-trials incorporating lead-in periods were excluded; SA3-studies with 'intensive' behaviour modification (BMOD and XENDOS) were excluded; SA4-trials with lead-in periods or 'intensive' behaviour modification were excluded.

ERG comment: The company did not actively search for trials comparing different types of behavioural interventions. Therefore, the CS includes only comparisons of NB32 plus standard management versus placebo plus standard management and NB32 plus intensive behaviour modification versus placebo plus intensive behaviour modification. There is no comparison of NB32 plus standard management versus intensive behaviour modification. In addition, the company did not include an analysis of NB32 plus intensive behaviour modification versus orlistat plus intensive behaviour modification. We have added this analysis in Section 4.5.2 of this report (using data from COR-BMOD for NB32 and XENDOS for orlistat).

4.4 Critique of the indirect comparison and/or multiple treatment comparison

Baseline characteristics for the four NB32 trials are reported in Table 4.5 in Section 4.2. Baseline characteristics for the 16 orlistat trials are reported in the table below.

One of the five NB32 trials and four out of the 16 orlistat trials included people with type 2 diabetes (T2DM) only. Participants in these trials were generally older and had more often hypertension and dyslipidaemia.

Table 4.21: Participant characteristics in the orlistat trials

	Astrup 2012[41]		Bakris 2002[42]		Berne 2005[43]		Broom 2002[44]	
	ORL	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**
n	95	98	267	265	111	109	259	263
Age, mean years ± SD	45.9 ± 9.1	45.9 ± 10.3	53.2 ± 0.5	52.5 ± 0.5	58.9 ± 9.1	59.3 ± 8.5	46.7 ± 11.4	45.3 ± 11.5
Age range, min-max	NR	NR	NR	NR	NR	NR	22-73	20-74
Female, n (%)	73 (77)	73 (75)	169 (63)	156 (59)	50 (45)	50 (46)	202 (78)	207 (79)
Weight, mean Kg ± SD	96.0 ± 1.7	97.3 ± 12.3	101.2 ± 1.0	101.5 ± 1.0	95.3 ± 12.6	95.7 ± 12.5	100.9 ± 20.5	101.8 ± 19.8
BMI, mean Kg/m^2 ± SD	34.1 ± 2.6	34.9 ± 2.8	35.8 ± 3.9	35.4 ± 4.0	32.6 ± 3.1	32.9 ± 3.0	37.1 ± 6.4	37.0 ± 6.2
Waist circumference, cm ± SD	108 ± 9.7	108 ± 10.0	108.6 ± 12.2	110.8 ± 12.5	108.0 ± 9.0	109.0 ± 9.3	107.8 ± 15.6	108.6 ± 16.4
White, n	NR	NR	CAU 226 HIS 10	228 CAU 3 HIS	NR	NR	NR	NR
Black, n	NR	NR	AF-AM 20	AF-AM 31	NR	NR	NR	NR
Asian, n	NR	NR	NR	NR	NR	NR	NR	NR
Other, n	NR	NR	1	2	NR	NR	NR	NR
T2DM, n (%)	3 (3)	4 (4)	23 (8)	22 (8)	111 (100)	109 (100)	IGT 11 (4)	IGT 15 (5)
Hypertension, n (%)	NR	NR	267 (100)	265 (100)	AHD 50 (45)	AHD 49 (45)	54 (20)	59 (22)
Dyslipidaemia, n (%)	NR	NR	LDL/HDL, mean ± SD 3.18 (1.1)	LDL/HDL, mean ± SD 3.3 (1.2)	LLD 13 (12)	LLD 17 (16)	114 (43)	120 (45)

Key: AF-AM = African-American, AHD = Participants taking antihypertensive drugs, CAU = Caucasian, HIS = Hispanics, IGT = Impaired Glucose Tolerance, LLD = Lipid Lowering Drugs, NR = Not Reported, ORL = Orlistat, PBO = Placebo, T2DM = Type 2 Diabetes Mellitus

Notes: No data were reported for BMI ranges, smokers, alcohol use, history of depression or anxiety

Table 4.21: Participant characteristics in the orlistat trials (continued)

	Derosa 2003*[45]		Derosa 2010[46]		Gotfredsen 2001[47]		Karhunen 2000[48]	
	ORL	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**
n	52	47	126	128	16	14	36	36
Age, mean years ± SD	52.3 ± NR	51.5 ± NR	53.0 ± 6	52.0 ± 5	42.2 ± 11.7	40.2 ± 9.6	42.9 ± NR	44.4 ± NR
Female, n (%)	26 (50)	25 (53)	62 (49)	64 (50)	13 (81)	13 (93)	NR	NR
Weight, mean Kg ± SD	95.1	95.3	94.5 ± 9.6	91.7 ± 8.7	107.6 ± 17.7	99.4 ± 9.2	98.1 ± 12.2	97.3 ± 14.8
BMI, mean Kg/m^2 ± SD	32.2	31.9	33.1 ± 2.9	32.5 ± 2.3	36.9 ± 3.9	36.6 ± 3.9	35.7 ± 3.4	36.1 ± 4.4
Waist circumference, cm ± SD	102.1	102.2	102. 0 ± 6.0	101.0 ± 5.5	NR	NR	106.8 ± 10.5	106.2 ± 11.2
T2DM, n (%)	NR	NR	126 (100)	128 (100)	NR	NR	0	0
Hypertension, n (%)	NR	NR	93 (86.1)	89 (80.2)	NR	NR	NR	NR
Dyslipidaemia, n (%)	100	100	23 (21.3)	21 (18.9)	NR	NR	NR	NR
Smoker, n (%)	NR	NR	41 (32.5)	46 (35.9)	NR	NR	NR	NR

Key: NR = Not Reported, ORL = Orlistat, PBO = Placebo, T2DM = Type 2 Diabetes Mellitus, * ORL 120mg TID and ORL120mg TID+FV have been pooled and PBO and FV arms have been pooled.

Notes: No data were reported for age range, ethnicity, BMI ranges, alcohol use, history of depression or anxiety

Table 4.21: Participant characteristics in the orlistat trials (continued)

	Kelley 2002[49]		Lindgarde 2000[50]		Lucas 2003[51]		Mathus-Vliegen 2006[52]	
	ORL	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**
n	266	269	190	186	256	188	14	14
Age, mean years ± SD	57.8 ± 8.1	58 ± 8.2	53.7 ± 9.4	53.2 ± 9.9	48.0 ± NR	48.0 ± NR	42.0 ± 11.7	45.5 ± 9.3
Age range, min-max	NR	NR	27-74	28-75	NR	NR	NR	NR
Female, n (%)	150 (56)	151 (56)	124 (65)	115 (62)	199 (78)	158 (84)	NR	NR
Weight, mean Kg ± SD	102.0 ± 1	101.8 ± 1	96.1 ± 13.7	95.9 ± 3.5	98.6	99.2	102.6 ± 12.3	109.3 ± 16.4
BMI, mean Kg/m^2 ± SD	35.8 ± 0.2	35.6 ± 0.3	33.2 ± 3.0	33.2 ± 3.1	35.7	36.2	35.7 ± 3.8	37.6 ± 3.9
Waist circumference, cm ± SD	113.1 ± 0.7	113.9 ± 0.8	106.0 ± 10.8	106.0 ± 11.0	NR	NR	NR	NR
T2DM, n (%)	100	100	54 (28.0)	44 (24.0)	NR	NR	0	0
Hypertension, n (%)	NR	NR	143 (82.0)	137 (74.0)	NR	NR	NR	NR
Dyslipidaemia, n (%)	NR	NR	HC 75 (39.0)	HC 75 (40.0)	100	100	NR	NR

Key: HC= hypercholesterolemia, NR = Not Reported, ORL = Orlistat, PBO = Placebo, T2DM = Type 2 Diabetes Mellitus
Notes: No data were reported for ethnicity, BMI ranges, smokers, alcohol use, history of depression or anxiety

Table 4.21: Participant characteristics in the orlistat trials (continued)

	Miles 2002[53]		Reaven 2001[54]		Swinburn 2005[55]		Torgerson 2004[56]	
	ORL	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**	**ORL**	**PBO**
n	250	254	156	91	170	169	1640	1637
Age, mean years ± SD	52.5 ± 0.4	53.7 ± 0.4	45.1 ± NR	44.1 ± NR	52.0 ± 7.5	52.5 ± 7.4	43.0 ± 8.0	43.7 ± 8.0
Female, n (%)	120 (48)	122 (48)	109 (70)	65 (71)	104 (61)	89 (52)	905 (55)	905 (55)
Weight, mean Kg ± SD	102.1 ± 1.0	101.1 ± 1.1	101.0 ± NR	101.1 ± NR	103.3 ± 17.8	106.9 ± 17.8	110.4 ± 16.3	110.6 ± 16.5
BMI, mean Kg/m^2 ± SD	35.6 ± 0.3	35.2 ± 0.2	35.6 ± NR	35.3 ± NR	37.6 ± 5.1	38.0 ± 4.9	37.3 ± 4.2	37.4 ± 4.5
Waist circumference, cm ± SD	NR	NR	NR	NR	112.4 ± 12.8	114.8 ± 13.1	115.0 ± 10.4	115.4 ± 10.4
White, n	CAU 211	CAU 201	NR	NR	NR	NR	NR	NR
Black, n	24	36	NR	NR	NR	NR	NR	NR
Asian, n	NR	NR	NR	NR	NR	NR	NR	NR
Other, n	15	17	NR	NR	NR	NR	NR	NR
T2DM, n (%)	100	100	0	0	14 (8)	14 (8)	0	0
Hypertension, n (%)	NR	NR	NR	NR	25 (15)	31 (18)	NR	NR
Dyslipidaemia, n (%)	NR	NR	NR	NR	HC 51 (30)	HC 49 (29)	NR	NR

Key: CAU = Caucasian, HC= hypercholesterolemia, NR = Not Reported, ORL = Orlistat, PBO = Placebo, T2DM = Type 2 Diabetes Mellitus
Notes: No data were reported for age range, ethnicity, BMI ranges, smokers, alcohol use, history of depression or anxiety

The orlistat trials were well conducted. Four trials were small (fewer than 100 patients) (Derosa 2003, Gotfredsen 2001, Karhunen 2000, Mathus-Vliegen 2006). The orlistat trials were older than the NB32 trials (2000 to 2012 with only two of 16 conducted in the last 10 years as opposed to 2010 to 2013 for the COR programme). Whilst the NB32 trials were conducted exclusively in the US, the orlistat trials were conducted across the world including the UK (Broom 2002), Italy (Derosa 2003 and Derosa 2010), The Netherlands (Mathus-Vliegen 2006), Sweden (Berne 2005, Lindgarde 2000, Torgersen 2004), Denmark (Gotfredsen 2001), Finland (Karhunen 2000), Europe-wide (Astrup 2012), the US (Bakris 2002, Kelley 2002, Lucas 2003, Reaven 2001), the US and Canada (Miles 2002) and Australia and New Zealand (Swinburn 2005).

Mean age of participants varied across the orlistat trials from 41 to 59 years. The NB32 trials ranged from 44 to 54 years for COR-DM. Where reported in the orlistat trials, percentages of female participants varied from 45% to 87%. Most trials had a reasonable proportion of male participants. In contrast 85% to 90% of participants in NB32 trials were female with only COR-DM recruiting 44% males. Only COR-DM recruited patients with diabetes but half of the orlistat trials included at least some patients with diabetes.

According to the CS, 11 orlistat trials had a lead-in period prior to randomisation in which no patients were excluded due to lack of efficacy or treatment compliance. A sensitivity analysis was performed excluding these trials. The NB32 trials did not have a lead-in period.

Across the orlistat trials and between the orlistat and NB32 trials there was variation in the components and delivery of standard care. Standard care is generally not reported in sufficient detail to assess comparability between trials.

Overall, the COR trials and orlistat trials appear comparable. The main difference appears to be that most of the orlistat trials have a more even gender balance than the NB32 trials which are conducted predominantly in women.

For the first outcome (mean percentage weight change from baseline at one year), the analyses performed are shown in Table 4.22, below.

Table 4.22: Number of studies reporting data for ≥5% reduction in weight at one year

Analysis	Trials with patients with T2DM only		Trials excluding patients with T2DM		All trials regardless of T2DM	
	NB32	ORL	NB32	ORL	NB32	ORL
Base case: All trials included	1	3	3	2	4	8
SA1: Trials with 'high' co-morbidities were excluded	0[a]	0[a]	3[b]	2[b]	3	3
SA2: Trials with lead-in periods were excluded	1	2	3	1	4	4
SA3: Trials with intensive BMOD were excluded	1[b]	3[b]	2	1	3	7
SA4: Trials with lead-in periods or intensive BMOD were excluded	1[c]	2[c]	2[a]	0[a]	3	3

Source: CS, Table 34, page 132

Notes: a, Insufficient data available to perform analysis; b, Analysis not performed as evidence base the same as the base case analysis; c, Analysis not performed as evidence base the same as SA2.

BMOD = behaviour modification; NB32 = naltrexone 32mg plus bupropion; NMA = network meta-analysis; SA = sensitivity analysis T2DM = Type 2 diabetes mellitus.

Results are presented as ORs with 95% CI for the direct meta-analyses and as ORs with 95% CrI for the Bayesian NMA (see Table 4.23). An OR < 1 favours NB32 over orlistat or placebo.

For patients with T2DM only, results of sensitivity analyses were either similar to the base case or not performed. Therefore, we will only present base case results.

For patients without T2DM, results from SA3 (excluding trials with intensive BMOD) were the only sensitivity analysis with results different from the base case analysis. Therefore, we will only present base case and SA3 results.

For all patients combined, results of sensitivity analyses were similar to the base case. Therefore, we will only present base case results.

As can be seen from Table 4.23, the Bayesian NMA found no significant differences between NB32 and orlistat for T2DM patients and for all patients combined. There is a statistically significant difference favouring NB32 over orlistat in the analyses excluding studies with T2DM patients, which indicates that more patients receiving NB32 had a ≥ 5% reduction in weight at one year compared to those receiving orlistat. The largest difference was seen in the third sensitivity analysis, where studies with ‘intensive’ behaviour modification (BMOD and XENDOS) were also excluded.

Table 4.23: Results for ≥5% reduction in weight at one year

	PLA	NB32	Placebo vs NB32 (OR, 95% CI)*	PLA	ORL	Placebo vs Orlistat (OR, 95% CI)*	Orlistat vs NB32 (OR, 95% CrI)**
T2DM only	30/159	118/265	0.29 (0.18, 0.46)	87/632	236/617	0.25 (0.17, 0.36)	FE1.09 (0.63, 1.88)
No T2DM - Base case - SA3	 244/1160 162/967	 901/1655 581/1173	 0.25 (0.17, 0.36) 0.21 (0.17, 0.25)	 765/1735 27/98	 1236/1735 42/95	 0.35 (0.23, 0.52) 0.48 (0.26, 0.87)	 FE 0.77 (0.61, 0.96) FE 0.44 (0.23, 0.84)
All patients	274/1319	1019/1920	0.26 (0.19, 0.34)	1052/3101	1841/3068	0.32 (0.26, 0.39)	RE 0.80 (0.51, 1.28)

*) Frequentist Odds Ratio (Non-event) (M-H, Random, 95% CI)
**) Bayesian NMA (OR, 95% CrI)
An OR < 1 favours the second treatment over the first. There are small differences with the results presented in CS because the company presented fixed effect results and we present random effects results for the direct meta-analysis .
RE = results from random effects NMA models which were presented for all patients, FE = results from fixed effect NMA models which were presented for the type 2 DM and no type 2 DM groups due to problems with Bayesian model convergence

For the second outcome (Mean percentage weight change from baseline at one year), the analyses performed are shown in Table 4.24, below.

Table 4.24: Number of studies reporting data for mean percentage weight CFB at one year

Analysis	Trials with patients with T2DM only		Trials excluding patients with T2DM		All trials regardless of T2DM	
	NB32	ORL	NB32	ORL	NB32	ORL
Base case: All trials included	1	4	3	5	4	16
SA1: Trials with 'high' comorbidities were excluded	0[a]	0[a]	3[b]	5[b]	3	8
SA2: Trials with lead-in periods were excluded	1	3	3	1	4	5
SA3: Trials with intensive BMOD were excluded	1[b]	4[b]	2	4	3	15
SA4: Trials with lead-in periods or intensive BMOD were excluded	1[c]	3[c]	2[a]	0[a]	3	4

Source: CS, Table 35, page 133

Notes: a, Insufficient data available to perform analysis; b, Analysis not performed as evidence base the same as the base case analysis; c, Analysis not performed as evidence base the same as SA2.

BMOD = behaviour modification; CFB = change from baseline; NB32 = naltrexone 32mg plus bupropion; NMA = network meta-analysis; SA = sensitivity analysis T2DM = Type 2 diabetes mellitus.

Results are presented as MDs with 95% CIs for the direct meta-analyses and as MDs with 95% CrIs for the Bayesian NMA (see Table 4.25). A MD > 0 favours NB32 over orlistat or placebo and indicates greater % weight reduction.

For patients with T2DM only, results of sensitivity analyses were either similar to the base case or not performed. Therefore, we will only present base case results.

For patients without T2DM, results from SA3 (excluding trials with intensive BMOD) were the only sensitivity analysis with results different from the base case analysis. Therefore, we will only present base case and SA3 results.

For all patients combined, results of sensitivity analyses were similar to the base case. Therefore, we will only present base case results.

As can be seen from Table 4.25, the Bayesian NMA found no significant differences between NB32 and orlistat for T2DM patients and for all patients combined. There is a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients are excluded. The largest difference was seen in the third sensitivity analysis, where studies with ‘intensive’ behaviour modification (BMOD and XENDOS) were also excluded.

Table 4.25: Results for mean percentage weight CFB at one year

	Placebo vs NB32 (MD, 95% CI)*	Placebo vs Orlistat (MD, 95% CI)*	Orlistat vs NB32 (MD, 95% CrI)**
T2DM only	3.20 (2.22, 4.18)	3.63 (2.30, 4.96)	FE 0.21 (-0.87, 1.30)
No T2DM - Base case - SA3	 4.88 (4.34, 5.42) 5.00 (4.41, 5.59)	 2.83 (1.41, 4.25) 2.01 (0.75, 3.27)	 FE 1.13 (0.44, 1.80) FE 2.98 (1.60, 4.36)
All patients	4.39 (3.49, 5.29)	3.00 (2.31, 3.69)	RE 1.39 (-0.08, 2.82)

*) Frequentist Mean Difference (IV, Random, 95% CI)
**) Bayesian NMA
A MD > 0 favours the second treatment over the first and indicates greater % weight reduction.
There are small differences with the results presented in CS because the company presented fixed effect results and we present random effects results for the direct meta-analyses.
RE = results from random effects NMA models which were presented for all patients, FE = results from fixed effect NMA models which were presented for the type 2 DM and no type 2 DM groups due to problems with Bayesian model convergence

ERG comment: Our main problem with these analyses is the use of the mITT populations for the NB32 trials, which we think produce biased results (see Section 4.2.3 of this report). Therefore, we have added analyses using true ITT populations from the NB32 trials in Section 4.5.1 of this report.

4.5 Additional work on clinical effectiveness undertaken by the ERG

In this section we will present two additional analyses:

- Results based on the ITT populations for the NB32 trials
- A comparison of studies with 'intensive' behaviour modification (BMOD vs. XENDOS)
- ERG preferred analyses, including full ITT data and no pooling of NB32 trials

4.5.1 Results based on the ITT populations for the NB32 trials

The company submission used modified intention-to-treat (mITT) data in their analyses. This is common in obesity trials. In fact, most orlistat trials used mITT data in their analyses, which usually included all randomised participants who had a valid baseline measurement and at least one valid measurement after randomisation. The definition of the mITT population in the NB32 trials is quite similar: 'all randomised patients with a post-baseline body weight measurement obtained while the patient remained on study medication'.

We agree that patients should have a baseline weight because otherwise there would be no possibility to calculate weight change. But including only patients that also have at least one post-baseline measurement can introduce bias, because the reason for missing post-baseline measurements could be related to the effectiveness of the treatment. Therefore, ideally the investigators should also present ITT results with some sensitivity analyses looking at different methods of imputing missing follow-up weights.

Additionally, the modified ITT population used in the NB32 trials is different from the mITT population used in the orlistat trials. The term mITT population is therefore misleading. This becomes clear when we look at the difference in the numbers of patients randomised and analysed in the trials. In the NB32 trials, 3,239 patients were analysed out of 3,958 randomised (81.8%); while in the orlistat trials 7,640 patients were analysed out of 7,754 randomised (98.5%). In the intervention arms this was 1,960 patients analysed out of 2,510 randomised (78.1%) for NB32 and 3,884 patients analysed out of 3,946 randomised (98.4%) for orlistat. In other words, in the NB32 trials, 21.9% of patients receiving NB32

were randomised but excluded from the analyses against 1.6% of patients receiving orlistat. Therefore, results of the mITT analyses in the orlistat are more or less the same as the ITT analyses; but in the NB32 trials there may be considerable differences between the types of analyses.

We will present an overview of results for the two main outcome measures (≥5% reduction in weight and percentage weight change from baseline) based on three different analyses: the mITT analysis as presented in the CS, and two ITT analyses (ITT-BOCF = all randomised patients with baseline-observation-carried-forward analysis; and ITT-Imp = all randomised patients with weight regain imputation method analysis). Results presented here are based on the same trial inputs as in the company submission. Therefore, any differences in results are a consequence of ITT versus mITT analyses. That means we have included all four NB32 trials for 'all patients' and three NB32 trials (COR-I, COR-II and COR-BMOD) for 'No type 2 DM'.

For the first outcome, Table 4.26 presents the data used from the four NB32 trials for each analysis and Table 4.27 presents the results for NB32 versus placebo and orlistat for each analysis and population.

Table 4.26: Data synthesised in analyses for ≥5% reduction in weight at one year

Study name	Arm 1	Arm 2	n1	r1	n2	r2
Greenway 2010 (COR-I)						
mITT	NB32	PBO	471	226	511	84
ITT-Imp	NB32	PBO	583	203	581	78
ITT-BOCF	NB32	PBO	583	180	581	67
Apovian 2013 (COR-II)						
mITT	NB32	PBO	702	355	456	78
ITT-Imp	NB32	PBO	878	337	495	73
ITT-BOCF	NB32	PBO	878	308	495	58
Hollander 2013 (COR-DM)						
mITT	NB32	PBO	265	118	159	30
ITT-Imp	NB32	PBO	335	104	170	27
ITT-BOCF	NB32	PBO	335	94	170	24
Wadden 2011 (COR-BMOD)						
mITT	NB32	PBO	482	320	193	82
ITT-Imp	NB32	PBO	591	NR	202	NR
ITT-BOCF	NB32	PBO	591	NR	202	NR
Sensitivity analysis*	NB32	PBO	565	321	196	84

BMOD = behaviour modification; COR = Contrave® obesity research; DM = diabetes mellitus; ITT-BOCF, all randomised patients with baseline-observation- carried-forward analysis; ITT-Imp, all randomised patients with weight regain imputation method analysis; mITT , modified intention-to-treat analysis; n = number of patients; NB32 = naltrexone 32mg plus bupropion; NR = not reported; r = number of patients achieving ≥5% reduction in weight; PBO = placebo.
*) Post-hoc sensitivity analysis for all randomised patients with a baseline and at least one post-baseline body weight measurement (see results on page 72 of CSR[35])

Table 4.27: Bayesian NMA results for ≥5% reduction in weight at one year

Population	Analysis	Orlistat vs NB32	Placebo vs NB32
All patients, mITT	RE	0.80 (0.51, 1.28)	0.25 (0.18, 0.37)
Type 2 DM, mITT	FE	1.09 (0.63, 1.88)	0.29 (0.18, 0.46)
No type 2 DM, mITT	FE	0.77 (0.61, 0.96)	0.24 (0.20, 0.29)
All patients, ITT-Imp	RE	1.14 (0.70, 1.91)	0.36 (0.25, 0.55)
Type 2 DM, ITT-Imp	FE	1.58 (0.91, 2.73)	0.41 (0.25, 0.66)
No type 2 DM, ITT-Imp	FE	1.09 (0.87, 1.36)	0.34 (0.29, 0.40)
All patients, ITT-BOCF	RE	1.11 (0.67, 1.91)	0.36 (0.24, 0.55)
Type 2 DM, ITT-BOCF	FE	1.59 (0.89, 2.79)	0.42 (0.25, 0.68)
No type 2 DM, ITT-BOCF	FE	1.06 (0.84, 1.33)	0.33 (0.28, 0.40)

Results are OR with 95% credible intervals (CrI). An OR < 1 favours the second treatment over the first.
Note: FE model results were presented for the Type 2 DM group and no type 2 DM groups due to problems with model convergence, RE model results were presented for all patients.
DM = diabetes mellitus; FE = fixed effect; ITT-BOCF = all randomised patients with baseline-observation-carried-forward analysis; ITT-Imp = all randomised patients with weight regain imputation method analysis; mITT = modified intention-to-treat analysis; NB32 = naltrexone 32mg plus bupropion; RE = random effects.

For the second outcome, Table 4.28 presents the data used from the four NB32 trials for each analysis and Table 4.29 presents the results for NB32 versus placebo and orlistat for each analysis and population.

Table 4.28: Data synthesised in analysis for percentage weight change from baseline at one year

Study name	Arm 1	Arm 2	n1	M1	SE1	n2	M2	SE2
Greenway 2010 (COR-I)								
mITT	NB32	PBO	471	-6.10	0.30	511	-1.30	0.30
ITT-Imp	NB32	PBO	583	-4.6	0.3	578	-1.2	0.3
ITT-BOCF	NB32	PBO	583	-4.0	0.3	578	-0.9	0.3
Apovian 2013 (COR-II)								
mITT	NB32	PBO	702	-6.40	0.30	456	-1.20	0.30
ITT-Imp	NB32	PBO	878	-4.9	6.5 (SD)	495	-1.2	6.7 (SD)
ITT-BOCF	NB32	PBO	878	-4.4	0.2	495	-0.8	0.3
Hollander 2013 (COR-DM)								
mITT	NB32	PBO	265	-5.00	0.30	159	-1.80	0.40
ITT-Imp	NB32	PBO	335	-3.5	0.3	170	-1.7	0.4
ITT-BOCF	NB32	PBO	335	-3.1	0.3	170	-1.3	0.4
Wadden 2011 (COR-BMOD)								
mITT	NB32	PBO	482	-9.30	0.40	193	-5.10	0.60
ITT-Imp	NB32	PBO	591	NR	NR	202	NR	NR
ITT-BOCF	NB32	PBO	591	-5.9	0.4	202	-4.0	0.6

BMOD = behaviour modification; COR = Contrave® obesity research; DM, diabetes mellitus; M = mean; n = number of patients; NB32 = naltrexone 32mg plus bupropion; NR = not reported; PBO = placebo; SD = standard deviation; SE = standard error.

Table 4.29: Bayesian NMA results for percentage weight change from baseline at one year

Population	Analysis	Orlistat vs NB32	Placebo vs NB32
All patients, mITT	RE	1.39 (-0.08, 2.82)	4.38 (3.15, 5.63)
Type 2 DM, mITT	FE	0.21 (-0.87, 1.30)	3.21 (2.23, 4.21)
No type 2 DM, mITT	FE	1.13 (0.44, 1.80)	4.88 (4.35, 5.43)
All patients, ITT-Imp	RE	0.26 (-1.23, 1.71)	3.25 (1.98, 4.51)
Type 2 DM, ITT-Imp	FE	-1.21 (-2.30, -0.11)	1.80 (0.83, 2.79)
No type 2 DM, ITT-Imp	FE	-0.09 (-0.77, 0.58)	3.65 (3.15, 4.17)
All patients, ITT-BOCF	RE	-0.31 (-1.81, 1.09)	2.68 (1.38, 3.89)
Type 2 DM, ITT-BOCF	FE	-1.21 (-2.30, -0.11)	1.80 (0.83, 2.79)
No type 2 DM, ITT-BOCF	FE	-0.54 (-1.21, 0.12)	3.20 (2.70, 3.71)

Results are mean difference with 95% credible intervals (CrI).
A MD > 0 favours the second treatment over the first and indicates greater % weight reduction.
Note: FE model results were presented for the Type 2 DM group and no type 2 DM groups due to problems with model convergence, RE model results were presented for all patients.
DM = diabetes mellitus; FE = fixed effect; ITT-BOCF = all randomised patients with baseline-observation-carried-forward analysis; ITT-Imp = all randomised patients with weight regain imputation method analysis; mITT = modified intention-to-treat analysis; NB32 = naltrexone 32mg plus bupropion; RE = random effects.

As can be seen from Tables 4.28 and 4.29, the positive effects of NB32 when compared to orlistat have all disappeared. For the first outcome (≥5% reduction in weight at one year), there was a statistically significant difference using mITT data favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded. In both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: OR = 1.09 (95% CrI: 0.87 to 1.36), ITT-BOCF: OR = 1.06 (95% CrI: 0.84 to 1.33). Moreover, although none of the differences are statistically significant, all results now favour orlistat.

For the second outcome (mean percentage weight change at one year), there was a statistically significant difference using mITT data favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded. In both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: MD= -0.09 (95% CrI: -0.77 to 0.58), ITT-BOCF: MD = -0.54 (95% CrI: -1.21 to 0.12). Moreover, although most of the differences are not statistically significant, most results now favour orlistat.

4.5.2 Comparison of intensive trials BMOD and XENDOS

One trial of orlistat[56] and one of NB32[26] were considered to include 'intensive' behaviour therapy. Brief details of the criteria used to define 'intensive' behaviour modification were provided in section 4.10.4 of the CS. However exact details of the criteria used were not provided. These two trials were excluded in sensitivity analysis 3 of the network meta-analysis to assess the robustness of the treatment effect.

XENDOS was a four year trial of orlistat conducted in Sweden. COR-BMOD was a 56 week trial of NB32 conducted in the US. XENDOS randomised 3,305 patients and COR-BMOD 793. A comparison of the participants, interventions and comparators, outcomes and study designs is given in Table 4.30.

Table 4.30: Comparison of intensive trials: COR-BMOD and XENDOS

	COR-BMOD	XENDOS
Participants	• Age 18 to 65 with • BMI 30 to 45 kg/m2 and uncomplicated obesity OR • BMI 27 to 45 kg/m2 and controlled hypertension and / or dyslipidaemia • Patients with diabetes excluded	• Age 30 to 60 with • BMI >= 30 • Patients with diabetes excluded
Intervention and Comparator	NB32+Behaviour modification(BMOD) vs. Placebo + BMOD BMOD consisted of group meetings lasting 90 minutes weekly for the first 16 weeks, then every other week for the next 12 weeks and monthly thereafter. They included instructions to consume a balanced deficit diet and to increase to 180 min/week of planned, moderately vigorous, physical activity (CS, page 57).	Orlistat +Lifestyle changes vs. Placebo + Lifestyle changes All patients prescribed a reduced calorie diet (approx. 800 kcal/day deficit) 30% from fat and no more than 300mg cholesterol per day. Readjusted every six months to account for weight loss during preceding months. Dietary counselling every 2 weeks for first 6 months and monthly thereafter. All kept physical activity diaries.
Primary outcome	Percentage of change in total body weight and proportion of patients with ≥ 5% decrease in total body weight at week 56 using modified intention-to-treat data.	Time to onset of type 2 diabetes and change in body weight after 4 years' treatment using intention-to-treat data.
Study design	RCT	RCT
Source: CS[1] and Torgerson 2004[56] (XENDOS)		

Details of the participant characteristics in the two trials can be found in Table 4.31.

Table 4.31: Comparison of participants in COR-BMOD and XENDOS

	COR-BMOD		XENDOS	
	NB32	Pbo	ORL	Pbo
No randomised	591	202	1640	1637
Age, mean years (SD)	45.9 (10.4)	45.6 (11.4)	43.0 (8.0)	43.7 (8.0)
Sex, female, n (%)	528 (89.3)	185 (91.6)	905 (55.2)	905 (55.3)
BMI, mean kg/m^2 (SD)	36.3 (4.2)	37.0 (4.2)	37.3 (4.2)	37.4 (4.5)
Weight, mean kg (SD)	100.2 (15.4)	101.9 (15.0)	110.4 (16.3)	110.6 (16.5)
Source: CS[1] and Torgerson 2004[56] (XENDOS)				

In XENDOS significantly more patients in the orlistat group (72.8%) than in the placebo group (45.1%) achieved weight loss ≥ 5% after one year of treatment. In BMOD 66.4% of patients in the NB32 group and 42.5% in the placebo group had a weight loss ≥ 5% (a statistically significant result, based on 482 and 193 patients in the mITT analysis). In terms of ≥ 10% weight loss, in XENDOS significantly more patients in the orlistat group were successful (41.0% of orlistat patients vs. 20.8% of placebo patients). In BMOD 41.5% of patients in the NB32 group and 20.2% in the placebo group had a weight loss ≥ 10% (a statistically significant result, based on 482 and 193 patients in the mITT analysis).

During the first year of treatment, the proportion of patients experiencing at least one gastrointestinal event with orlistat or placebo in XENDOS was 91% vs. 65%, respectively. In COR-BMOD 65.1% of patients experienced gastrointestinal disorders in the NB32 group and 39% in the placebo group. Overall, 4% of placebo patients and 8% of orlistat patients withdrew from XENDOS because of adverse events or laboratory abnormalities; the difference was primarily due to gastrointestinal events. In COR-BMOD a greater percentage of those in the NB32 group discontinued due to an adverse event (25.4% vs. 12.4%, $p < 0.001$).

ERG comment:

- Although interventions in COR-BMOD and XENDOS could both be considered intensive, the nature of the co-intervention delivered varied in terms of delivery, intensity and advice components.
- Participant inclusion criteria were similar and both trials excluded patients with diabetes. COR-BMOD had a greater proportion of female participants than XENDOS. Participants in XENDOS were, on average approximately 10kg heavier.
- XENDOS specifically considered time to onset of type 2 diabetes in addition to change in body weight as a primary outcome.
- Both trials found active treatment with a drug to be superior to lifestyle management alone in terms of 5% or 10% weight loss. Although there were a large number of gastrointestinal events in the XENDOS trial, discontinuation due to adverse events was lower than that noted in the COR-BMOD trial.

The results of the indirect comparison of NB32 plus intensive behaviour modification versus orlistat plus intensive behaviour modification, using data from COR-BMOD versus XENDOS, are presented in the Table below.

Table 4.32: Indirect comparison results for COR-BMOD versus XENDOS

Population	NB 32 vs placebo	Orlistat vs placebo	Orlistat vs NB32
≥5% reduction in weight at 1 year			
mITT	2.67 (1.90, 3.77)	3.26 (2.82, 3.77)	1.22 (0.84, 1.77)
ITT-Imp	NR		NR
ITT-BOCF	1.75 (1.26, 2.44)	3.26 (2.82, 3.77)	1.86 (1.30, 2.66)
Mean % weight CFB at 1 year			
mITT	-4.20 (-5.62, -2.78)	-3.99 (-4.46, -3.52)	-0.21 (-1.28, 1.70)
ITT-Imp	NR		NR
ITT-BOCF	-1.9 (-3.27, -0.53)	-3.99 (-4.46, -3.52)	-2.09 (-3.53, -0.65)

Results are OR with 95% CI for ≥5% reduction in weight at 1 year and mean difference (MD) with 95% CI for mean % weight CFB at 1 year. The analysis uses the Bucher method for indirect comparisons.
An OR < 1 favours the second treatment over the first. A MD > 0 favours the second treatment over the first and indicates greater % weight reduction.
CFB = Change from baseline; ITT-BOCF = all randomised patients with baseline-observation-carried-forward analysis; ITT-Imp = all randomised patients with weight regain imputation method analysis; mITT = modified intention-to-treat analysis; NB32 = naltrexone 32mg plus bupropion; NR = Not reported.

Results in Table 4.32 show that results are different dependent on which dataset is used. When using the mITT results for ≥ 5% reduction in weight at one year there is no significant difference between NB32 and orlistat. However when using the ITT BOCF results (ITT-Imp results were not available for

COR-BMOD) the results are statistically significant and favour orlistat (OR 1.86, 95% CI 1.30 to 2.66). For the percentage weight change from baseline there was also no significant difference between NB32 and orlistat when using the mITT results. However, when using the ITT-BOCF results there was a statistically significant difference which favoured orlistat (MD -2.09, 95% CI -3.53 to -0.65).

4.5.3 ERG preferred analyses

In Section 4.5.1 we have used the same trial inputs as in the company submission. Therefore, any differences in results were a consequence of ITT versus mITT analyses. That means we have included all four NB32 trials for 'all patients' and three NB32 trials (COR-I, COR-II and COR-BMOD) for 'No type 2 DM'.

As explained in Section 4.2.5, we would prefer not to pool any of the NB32 trials. That means, we will not present any results for 'all patients', because this would be a mix of diabetes patients in the COR-DM trial and non-diabetes patients in the other three COR trials. In addition, different interventions (standard management and intensive behaviour modification) would be pooled.

Therefore, we would preferably use only trials that include T2DM patients for obese patients with diabetes, and only trials that do not included T2DM patients for obese patients without diabetes in our analyses. In addition, we will not include trials with intensive behaviour modification (COR-BMOD for NB32, and XENDOS for orlistat) in these analyses. That means that the analyses for obese patients with diabetes are the same as before, but results for obese patients without diabetes will change as now only COR-I for NB32 and Astrup et al. (2012)[41] for orlistat are included.

Table 4.33 shows that the results for 'obese patients with T2DM' and 'intensive behaviour modification' are the same as in Sections 4.5.1 and 4.5.2, respectively. However, results for 'obese patients without T2DM' have changed considerably again, and are almost the same as in the company's original analyses. Both outcomes show no significant difference between NB32 and orlistat, but both favour NB32.

Table 4.33: ERG preferred analyses compared to other results

Population		Company analyses (mITT data)*	Company analyses (ITT-BCFA data)**	ERG preferred analyses**
		Orlistat vs NB32	Orlistat vs NB32	Orlistat vs NB32
Obese people with T2DM				
≥5% reduction in weight at 1 year	OR	1.09 (0.63 to 1.88)	1.59 (0.89 to 2.79)	1.59 (0.89 to 2.79)
Mean % weight CFB at 1 year	MD	0.21 (-0.87 to 1.30)	-1.21 (-2.30 to -0.11)	-1.21 (-2.30 to -0.11)
Obese people without T2DM				
≥5% reduction in weight at 1 year	OR	0.77 (0.61 to 0.96)	1.06 (0.84 to 1.33)	0.61 (0.31 to 1.22)
Mean % weight CFB at 1 year	MD	1.13 (0.44 to 1.80)	-0.54 (-1.21 to 0.12)	1.11 (-0.39 to 2.63)
Intensive behaviour modification				
≥5% reduction in weight at 1 year	OR	1.22 (0.84 to 1.77)	1.86 (1.30 to 2.66)	1.86 (1.30 to 2.66)
Mean % weight CFB at 1 year	MD	-0.21 (-1.28 to 1.70)	-2.09 (-3.53to -0.65)	-2.09 (-3.53to -0.65)

Population		Company analyses (mITT data)*	Company analyses (ITT-BCFA data)**	ERG preferred analyses**
		Orlistat vs NB32	Orlistat vs NB32	Orlistat vs NB32
Results are OR with 95% CI/CrI for ≥5% reduction in weight at 1 year and mean difference (MD) with 95% CI/CrI for mean % weight CFB at 1 year. An OR less than one favours NB32 over orlistat and a CI including 1 is not significant. A MD of >0 favours NB32 over orlistat and indicates greater % weight reduction and a CI including 0 is not significant. *) Bayesian NMA (OR, 95% CrI) using mITT data; **) Using the Bucher method for indirect comparisons and ITT-BCFA data. FE = fixed effect; ITT-BCFA = all randomised patients with baseline-carried-forward analysis; MD = Mean Difference; mITT = modified intention-to-treat analysis; NB32 = naltrexone 32mg plus bupropion; OR = Odds Ratio; T2DM = Type 2 diabetes mellitus;				

4.6 Conclusions of the clinical effectiveness section

The company conducted a systematic review to identify studies comparing NB32 to the comparators outlined in the NICE scope.[7] Relevant direct evidence comparing NB32 and placebo has been presented. However no trials directly comparing NB32 to orlistat were identified. Indirect comparisons were made between NB32 and orlistat.

The company submission focused on data from the four pivotal RCTs: COR-I, COR-II, COR-BMOD, and COR-DM. All of these RCTs compare NB32 to placebo with both arms receiving standard care. Standard care varies between the trials in that COR-BMOD has a more intensive form of behavioural management. In addition, COR-DM focused exclusively on patients with diabetes whilst the other trials exclude patients with diabetes. The ERG agrees that there was clinical and statistical heterogeneity between the four COR trials and that because of this the results from the separate analyses for patients with and without diabetes should be preferred and BMOD may not be suitable to be pooled with the other COR trials.

The NB-CVOT study was included in the submission as a supporting study as it presented longer term outcomes. NB-CVOT represents an older population with cardiovascular disease when compared to the COR trials. Most of the patients in NB-CVOT are diabetic, and many are depressed. A number of problems with the study were identified. NB-CVOT used a lead-in period where large numbers of patients discontinued primarily due to adverse events. This implies that those continuing to the treatment period who were re-randomised were better able to tolerate the drug. The adverse event profile will be an overestimate of the tolerability of the drug. In NB-CVOT only SAEs and AEs leading to study drug discontinuation were collected. Even so, an elevated number of gastrointestinal events were noted in the NB32 group. NB-CVOT was terminated early (after the 50% interim analysis), after 25% interim data were made public. The trial was not able to provide a definitive answer to the cardiovascular risk of NB32 and a further trial has been instigated. The reliability of the final data on weight loss is also questionable.

The COR trials were of high quality. However more patients dropped out of NB32 groups due to adverse events. Higher rates of adverse events (especially nausea) could have resulted in un-blinding of participants. The modified intention-to-treat analysis presented in the submission reflects only those who have a post-baseline measurement whilst on the study drug. Any discontinuations before the post-baseline weight assessment are discounted. Reasons for discontinuation could relate to efficacy or safety of the drug. Using the true ITT data, NB32 is still superior to placebo in terms of weight loss but results are more modest.

A number of points should be borne in mind when applying the results of the NB32 trials to clinical practice:

- Overweight patients in addition to obese patients were included in the NICE scope. However only a very small percentage (approximately 2%) of patients who are overweight are in the COR trials. Therefore this population is not well represented. Mean BMI in the trials is 36 to 37 which is severely obese.
- All of the COR trials were conducted in the US so participant characteristics and the nature of standard care may differ from a UK setting.
- Prior use of orlistat was an exclusion criterion in all four COR trials. Therefore the effect of NB32 on those who have failed on orlistat has not been examined.
- The majority of participants in the COR trials are female. The ERG draws to the attention of the committee that this does not reflect the distribution of obesity according to gender. Men in England are more likely to be overweight or obese (68% vs 58% in 2015).
- Asian patients are not well represented in the COR trials so results may not be applicable to these ethnic groups.
- Three of the four COR trials measure the primary outcome at 56 weeks. Although this is acceptable in terms of weight loss, there is no information on maintenance of weight loss after this time.
- The CS states that "*For patients continuing treatment post 16 weeks, treatment should be continued as long as clinical benefit is observed*."[1] It is unclear how long patients would continue to take the drug in practice.
- There are no data on the effectiveness of retreatment with NB32 following successful treatment with NB32 and subsequent discontinuation and weight gain.
- There were large dropout rates across the COR trials (up to 50%). This suggests that in practice up to half of patients may complete a year's treatment with NB32 which is relevant when considering transferability to clinical practice.
- Based on the mITT data presented by the company NB32 results in greater weight loss and in a higher number reporting 5% or more weight loss. However the superior results regardless of arm in the BMOD trial are of interest. NB32 together with a more intensive behaviour modification programme resulted in 66.4% of patients losing 5% or more weight compared to 44 to 55% in the other three trials without such an intensive intervention. In the BMOD trial the placebo and behaviour modification arm achieved results approaching the medication arms in the other trials.
- A greater proportion of gastrointestinal events, particularly nausea, were noted in NB32 groups across the trials. Although the majority of events were not serious, more participants withdrew as a result of adverse events in treatment groups.

A comparison between NB32 (plus standard management) versus intensive behaviour modification is missing. In its response to the clarification letter (Question A12, page 13), the company stated that "*the anticipated positioning of NB32 in the treatment pathway is for patients eligible for pharmacological treatment (alongside standard management)"*. Therefore, the company considered different types of behaviour modification not relevant to the decision problem. However, the NICE scope clearly mentions 'standard management without naltrexone-bupropion' as a comparator and this may very well include more intensive forms of behaviour modification than patients receiving instructions to follow a diet and increase physical activity, and written behaviour modification advice. Moreover, results from the COR-BMOD trial show that more intensive behaviour modification is still quite effective in patients eligible for pharmacological treatment. At first glance it seems that intensive behaviour modification in

the COR-BMOD trial (percentage change from baseline: -5.1 (SE: 0.6)) has similar effects as NB32 in the COR-I trial ((percentage change from baseline: -6.1 (SE: 0.3)). Therefore, a comparison of NB32 vs. intensive behaviour modification would have been of interest.

Regarding the comparison of NB32 with orlistat, the company used modified ITT data from NB32 trials, but this is misleading. The mITT population in the NB32 trials is very different from mITT populations in the orlistat trials. In the NB32 trials, 21.9% of patients receiving NB32 were randomised but excluded from the analyses against 1.6% of patients receiving orlistat.

The comparison with orlistat may be biased in favour of NB32. NB32 trials were published in 2010 or later; most of the trials with orlistat were published before 2005, so caution should be exercised when making indirect comparisons; this is particularly true for conditions such as diabetes where background standard therapy (for glucose and lipids especially) may be very different now.

We have reproduced the company's indirect analyses using full ITT data from the NB32 trials and we have included a new analysis: an indirect comparison of NB32 plus intensive behaviour modification (COR-BMOD) versus orlistat plus intensive behaviour modification (XENDOS). The results show that the positive effects of NB32 when compared to orlistat have all disappeared. For the first outcome (≥5% reduction in weight at one year), there was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded when using mITT data. In both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: OR = 1.09 (95% CrI: 0.87 to 1.36), ITT-BOCF: OR = 1.06 (95% CrI: 0.84 to 1.33). Moreover, although none of the differences are statistically significant, all results now favour orlistat.

For the second outcome (mean percentage weight change at one year), there was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded when using mITT data. In both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: MD= -0.09 (95% CrI: -0.77 to 0.58), ITT-BOCF: MD = -0.54 (95% CrI: -1.21 to 0.12). Moreover, although most of the differences are not statistically significant, most results now favour orlistat.

The results of the indirect comparison of NB32 plus intensive behaviour modification versus orlistat plus intensive behaviour modification, using data from COR-BMOD versus XENDOS, show that both outcomes significantly favour orlistat over NB32 (≥5% reduction in weight at one year: OR 1.86 (95% CI: 1.30 to 2.66); Mean percentage weight CFB at one year: MD -2.09 (95% CI: -3.53 to -0.65)). This is particularly relevant, as the committee might assume that those who are prescribed NB32 or orlistat might want to participate in a weight loss programme. In that case, the BMOD trial might provide a better estimate of the effect of NB32 as an adjunct to standard management.

Finally, we performed our preferred analyses, i.e. using full ITT data and no pooling of NB32 trials. The results for 'obese patients with T2DM' and 'intensive behaviour modification' are the same as before, but results for 'obese patients without T2DM' have changed considerably again, and are almost the same as in the company's original analyses. Both outcomes show no significant difference between NB32 and orlistat, but both favour NB32.

The table below shows the main results for obese people with diabetes, obese people without diabetes and NB32 plus intensive behaviour modification versus orlistat plus intensive behaviour modification.

Table 4.34: Company results versus ERG results

Population		Company analyses (mITT data)*	Company analyses (ITT-BCFA data)**	ERG preferred analyses**
		Orlistat vs NB32	Orlistat vs NB32	Orlistat vs NB32
Obese people with T2DM				
≥5% reduction in weight at 1 year	OR	1.09 (0.63 to 1.88)	1.59 (0.89 to 2.79)	1.59 (0.89 to 2.79)
Mean % weight CFB at 1 year	MD	0.21 (-0.87 to 1.30)	-1.21 (-2.30 to -0.11)	-1.21 (-2.30 to -0.11)
Obese people without T2DM				
≥5% reduction in weight at 1 year	OR	0.77 (0.61 to 0.96)	1.06 (0.84 to 1.33)	0.61 (0.31 to 1.22)
Mean % weight CFB at 1 year	MD	1.13 (0.44 to 1.80)	-0.54 (-1.21 to 0.12)	1.11 (-0.39 to 2.63)
Intensive behaviour modification				
≥5% reduction in weight at 1 year	OR	1.22 (0.84 to 1.77)	1.86 (1.30 to 2.66)	1.86 (1.30 to 2.66)
Mean % weight CFB at 1 year	MD	-0.21 (-1.28 to 1.70)	-2.09 (-3.53to -0.65)	-2.09 (-3.53to -0.65)

Results are OR with 95% CI/CrI for ≥5% reduction in weight at 1 year and mean difference (MD) with 95% CI/CrI for mean % weight CFB at 1 year.
An OR less than one favours NB32 over orlistat and a CI including 1 is not significant. A MD of >0 favours NB32 over orlistat and indicates greater % weight reduction and a CI including 0 is not significant.
*) Bayesian NMA (OR, 95% CrI) using mITT data; **) Using the Bucher method for indirect comparisons and ITT-BCFA data.
FE = fixed effect; ITT-BCFA = all randomised patients with baseline-carried-forward analysis; MD = Mean Difference; mITT = modified intention-to-treat analysis; NB32 = naltrexone 32mg plus bupropion; OR = Odds Ratio; T2DM = Type 2 diabetes mellitus;

5. COST EFFECTIVENESS

5.1 ERG comment on company's review of cost effectiveness evidence

This section pertains mainly to the review of cost effectiveness analysis studies. However, the search section (5.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the company submission. Therefore, the following section includes searches for the cost effectiveness analysis review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

5.1.1 Objectives of cost effectiveness searches and reviews

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the company submission.

Objectives of cost effectiveness analysis search and review

The CS reported that searches were carried out in May 2016. Searches contained a 10 year date limit, but were not limited by language. Searches were carried out on the following databases: Embase, MEDLINE, MEDLINE in-Process, HTA and NHS EED via the Cochrane library and Econlit. Searches were carried out in line with the NICE 2013 guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4.[57] Supplementary searches of the following conference proceedings were reported: International Congress on Obesity (ICO), European Congress on Obesity by the European Association for the Study of Obesity (ECO), International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Annual European Congress and ISPOR Annual International Congress. Along with searches of both the NICE and Public Health England websites, the CS also reported that "*bibliographic searches of published systematic reviews, economic models and health technology assessments (HTAs) were conducted.*"[38]

ERG comment: The majority of searches in Appendix 15 were well reported and easily reproducible. Table 22 reported the use of the SIGN study design filter for economics.[38] Unlike the clinical effectiveness section the filter devised for MEDLINE was used for the joint MEDLINE/Embase search, however the remaining condition and interventions facets of the strategy employed only Emtree terms and no MeSH. As stated in Section 4.1.1, although some mapping between indexing terms does take place on Embase.com it is possible that in this case some relevant Emtree/MeSH indexing terms will not be included in the search, and potentially relevant records could have been missed.

The ERG also noticed a number of areas for concern relating to the Econlit search reported for this section. Firstly the ERG asked the company to confirm that this search was conducted using the EBSCO platform as stated in Table 22, Appendix 15. The ERG noted the inclusion of the MH field tag in Table 26, which the ERG understands is not supported by EBSCO as field in Econlit. The company responded that the MH search functionality was incorrectly presented in Table 22. Further to this the ERG noted that the strategy appeared to contain an error in the line numbers being combined in lines S60 and S61. The line above (S59) had the combination "S11 AND S25 AND S58" which appeared to be correct; however the following two lines had the combination "S11 AND S22 AND S58. Line S25 was a combination of all listed interventions where line S22 was for "TI (lorcaserin OR belviq) OR AB (lorcaserin OR belviq)". The company confirmed that this was also due to a reporting error and provided a full revised strategy in their response to clarification. Finally despite being a pre-filtered specialist resource, as with the Cochrane strategies reported in the clinical effectiveness section, the Econlit strategy contained a redundant economics filter, which may have unnecessarily restricted the results

retrieved. However given other searches reported for this section, this is unlikely to have impacted on the overall recall of results.

Objectives of search and review for measurement and valuation of health effects

Searches were conducted to "*identify utility values associated with overweight (with at least one comorbidity) and obese conditions and their associated treatments*". [38]

Searches were carried out in June 2016 across a good range of databases. No date or language limits were applied. The company reported that the supplementary database and conference websites searched for modelling studies were also searched for utility studies.

ERG comment: Searches were well reported and easily reproducible. As with the previous sections the ERG had some concerns regarding the use of only Emtree indexing terms. Despite some mapping between indexing terms on Embase.com the same limitations as described in Section 4.1.1 will apply. The Econlit strategy for this section also contained the unsupported use of the MH field tag which the company reported as a presentation error in their response to clarification.

Objectives of search and review for cost and healthcare resource identification, measurement and valuation

A systematic literature review was conducted to "*identify the economic burden of obesity and associated treatments, in terms of healthcare resource utilisation as well as direct and indirect costs.*"[38]

Searches were carried out in June 2016 on a good range of databases. As with the previous sections supplementary searches of conference proceedings and other relevant websites were carried out in order to identify cost and resource use studies. As with the economics section, searches were limited to the last 10 years and for this section only data from the UK was sought.

ERG comment: Searches were well reported and easily reproducible. The same errors regarding the use of the unsupported MH field tag in the Econlit search and limited indexing terms on Embase.com searches appeared in these searches as for earlier sections, with regard to the latter the same limitations will apply.

5.1.2 Inclusion/exclusion criteria used in the study selection

The pre-specified eligibility are shown as a PICOS table in Table 52 of the CS[1] for cost effectiveness analysis studies, in Table 59 of the CS[1] for measurement and valuation of health effects studies, and in Table 61 of the CS[1] for cost and healthcare resource use studies.

ERG comment: The ERG was satisfied that the company's inclusion and exclusion criteria used in the study selection were appropriate for the three searches.

5.1.3 Included/excluded studies in the cost effectiveness review

The search identified 1,792 citations, of which 1,781 were identified through database searching, one additional study through bibliographic searching and 10 abstracts were identified from conference proceedings. After screening and eligibility assessment, 81 references were deemed eligible for full-text evaluation. Nineteen studies from 22 included publications met the inclusion criteria. Table 27 in Appendix 15 of the CS[1] provides a tabular overview of the included studies.

The following is an overview of the company's findings from the review, as reported in the CS[1]:

- None of the included studies considered NB32 as an intervention

- Four studies were set in the UK (Ara et al., 2012[58], Davies et al., 2012[59], Burch et al., 2009[60], Beaudet et al., 2011[61])
- Pharmacological treatment for obesity has the potential of being cost effective
- Results were particularly sensitive to uncertainty surrounding assumptions concerning duration of weight maintenance after initial weight loss and the effect of a reduction in BMI on health-related quality of life (HRQoL)
- A variety of model types and structures was used across the included studies, with most studies using timed cohort models, and one study using an individual-level timed model,[58] which was deemed most appropriate by the company.

The NICE and PHE website search identified only one result: the "Weight Management Economic Assessment Tool". The company reports that it "*helps healthcare professionals assess existing or planned weight management interventions and to allow commissioners to compare the costs of an intervention for English patients with potential cost savings*.[62]" The tool has been developed by PHE in conjunction with a panel of experts.[62]

ERG comment: The ERG considered that the searches and review were unlikely to miss any important studies and considers the company's conclusions as appropriate.

5.1.4 Conclusions of the cost effectiveness review

The CS provides an overview of the included studies but no specific conclusion regarding the cost effectiveness of NB32, or other pharmacological treatments, is formulated.

ERG comment: Since the identified studies did not consider NB32 as an intervention, the ERG agrees that no specific conclusion could be drawn from this review.

5.2 Summary and critique of company's submitted economic evaluation by the ERG

Table 5.1: Summary of the company's economic evaluation (with signposts to CS)

	Approach	Source / Justification	Signpost (location in CS)
Model	A DES model was implemented in Excel using the "discretely integrated condition event" (DICE) principles and structure	It was argued that an individual-level approach is better suited than a cohort-level approach to capture the chronic implications of both weight and weight-related health events in a heterogeneous group of overweight and obese patients.	Sections 5.2.2.1 and 5.2.2.2
States and events	Events: - treatment discontinuation - development of T2DM - first cardiovascular event - second cardiovascular event - death	The company used the economic evaluation by Ara et al.[58] as a starting point.	Sections 5.2.2.1 and 5.2.2.3
Comparators	- orlistat as an adjunct to standard management and;	Consistent with the scope and licensed indications	Section 5.2.4

	Approach	Source / Justification	Signpost (location in CS)
	- standard management alone		
Population	The company stated that the model aimed to reflect adult patients who are obese (BMI ≥30kg/m2) or overweight (BMI ≥27kg/m2 and <30kg/m2) in the presence of one or more weight-related comorbidities (e.g., T2DM, dyslipidaemia, or controlled hypertension).	NB32 is licensed as an adjunct to standard non-pharmacological management for this population.	Section 5.2.1
Treatment effectiveness	Treatment effectiveness is estimated based on reduced weight / BMI (retrieved from COR trial programme) and subsequent reduced risk of obesity-related comorbidities (based on the economic evaluation by Ara et al.[58]).		Sections 5.3.2 to 5.3.4
Adverse events	Costs were considered for AEs that occurred in at least 5% of patients (either treatment arm) in the COR-I trial. No disutilities related to adverse events were considered.	The 5% threshold was selected to reflect the British National Formulary criteria of all very common (> 1 in 10) and the majority of common (1 in 100 to 1 in 10) AEs. Moreover, the company stated that quality of life implications of adverse events were deemed to be too poorly understood to incorporate disutilities associated with adverse events.	Sections 5.4.3, 5.4.4 and 5.5.4
Health related QoL	The HRQL data used in the cost-effectiveness analysis are estimated based on Tobit regression analysis of EQ-5D individual-level data from a recent Health Survey for England.	The Tobit model includes explanatory variables for BMI, age, gender, and the obesity-related conditions in the economic model as well as cancer, and are therefore well suited to inform utility assumptions in the model.	Section 5.4
Resource utilisation and costs	Costs in the model consisted of drug acquisition costs, non-drug costs related to standard management (applicable to all	Considering the studies identified in the review, the company stated that the level of reporting was generally poor across studies, to the extent that it was difficult to	Section 5.5

	Approach	Source / Justification	Signpost (location in CS)
	treatments considered), obesity-related comorbidity costs and adverse event costs. These costs were primarily based on Ara et al.,[58] NHS reference costs and PSSRU.	elicit useful resource use estimates for this analysis. A notable exception to this was the study by Ara et al..[58] Hence the company used this study to inform healthcare resource use assumptions.	
Discount rates	Discount rate of 3.5% for utilities and costs	As per NICE reference case	Table 54
Sub groups	Stratified based on T2DM.	As per NICE scope	Section 5.9
Sensitivity analysis	Both DSA and PSA were performed as well as scenario analyses		Section 5.8
Source: CS Abbreviations: DSA, deterministic sensitivity analysis; PSA, probabilistic sensitivity analysis; PSSRU, Personal Social Services Research Unit; T2DM, type 2 diabetes; DES, discrete event simulation; BMI, body mass			

5.2.1 NICE reference case checklist (TABLE ONLY)

Table 5.2: NICE reference case checklist

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Population	As per NICE scope	Yes	
Comparator(s)	Therapies routinely used in the National Health Service (NHS), including technologies regarded as current best practice	Yes	
Type of economic evaluation	Cost effectiveness analysis	Yes	
Perspective on costs	NHS and Personal Social Services (PSS)	Yes	
Perspective on outcomes	All health effects on individuals	Yes	
Time horizon	Sufficient to capture differences in costs and outcomes	Yes	
Synthesis of evidence in outcomes	Systematic review	Yes	

Elements of the economic evaluation	Reference Case	Included in submission	Comment on whether *de novo* evaluation meets requirements of NICE reference case
Measure of health effects	Quality adjusted life years (QALYs)	Yes	
Source of data for measurement HRQoL	Described using a standardised and validated instrument	Yes	
Source of preference data for valuation of changes in HRQoL	Time-trade off or standard gamble	Yes	
Discount rate	An annual rate of 3.5% on both costs and health effects	Yes	
Equity weighting	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes	
Sensitivity analysis	Probabilistic modelling	Partly	The number of simulated patients (1,000) is too low to provide stable results The PSA does not incorporate all relevant parameters (e.g. the uncertainty surrounding time to treatment discontinuation, a key parameter in the model, was neglected). The number of PSA simulations (100) is too low to provide stable results.
Source: CS Abbreviation: PSA, probabilistic sensitivity analysis			

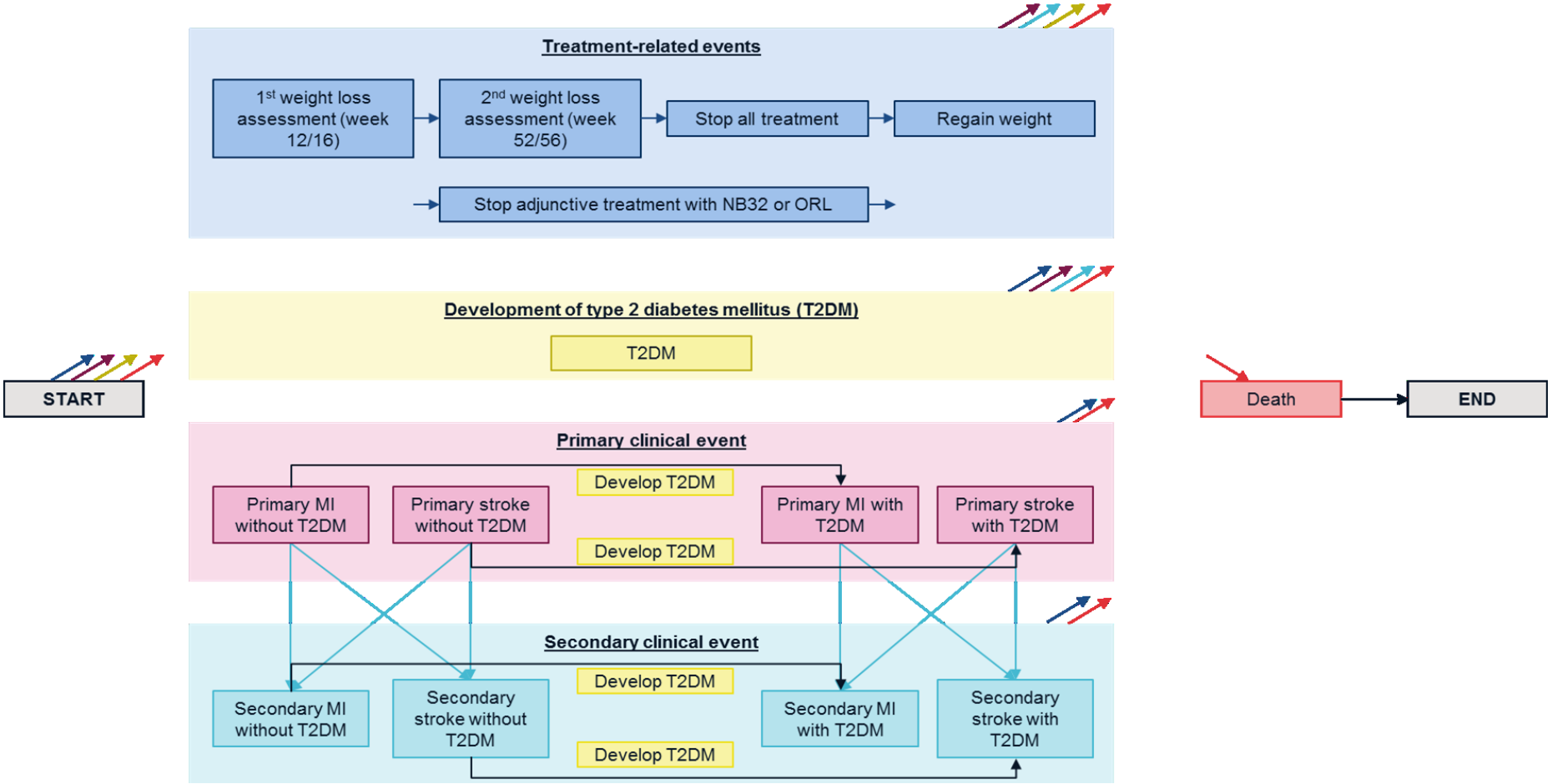
5.2.2 Model structure

The company developed a de novo economic model using an individual-level approach, more specifically a discrete event simulation (DES). It was argued that an individual-level approach is better suited than a cohort-level approach to capture the chronic implications of both weight and weight-related health events in a heterogeneous group of overweight and obese patients. The DES model was implemented in Excel using the "discretely integrated condition event" (DICE) principles and structure proposed by Caro.[63] In addition, the company used the economic evaluation by Ara et al.[58] (also an individual-level model) as a starting point, which is a Health Technology Appraisal report (2012) comparing different pharmacological treatments for obesity.

An overview of the de novo model structure is shown in Figure 5.1. This Figure aims to describe the logic and assumptions underpinning the model, by depicting the process of a simulated individual's progress through the model, from model entry ("START"), through the various treatment and disease

events that may occur in the model and have consequences for patient utility and/or health and social care costs, to death and model exit (“END”).

Figure 5.1: Model structure



Source: CS Figure 25
Abbreviations: MI, myocardial infarction, NB32, naltrexone 32mg plus bupropion; ORL, orlistat; T2DM, Type 2 diabetes mellitus.
Notes: Arrows demonstrate the possible transitions to each type of event.

As depicted in Figure 5.1, the following events are considered in the economic model:

- treatment discontinuation (either based on time to treatment discontinuation (TTD) for NB32 and orlistat, or based on weight loss assessment; see Sections 5.2.4 and 5.2.6 for more details);
- development of T2DM (based on model by Ara et al.[58]);
- first cardiovascular event (either stroke or MI, based on model by Ara et al.[58]);
- second cardiovascular event (either stroke or MI, based on model by Ara et al.[58]) and;
- death (first 15 years in the model based on model by Ara et al.[58]; afterwards based on general population mortality estimates from the Office for National Statistics National Life Tables).

Upon model entry a simulated patient is assigned a profile of sampled baseline characteristics that are explanatory factors for risks, costs and/or utility in the model (sampled baseline characteristics as well as random numbers for the sampled patient are equal across all three treatments). The baseline profile characterises the individual patients by:

- age (years);
- gender (male, female);
- height (meters);
- BMI (kg/m^2);
- T2DM status (yes, no);
- smoker status (current, previous, never);
- receive insulin, if diabetic (yes, no);
- receive statins (yes, no);

In addition to the characteristics listed above, history of angina, diabetes mellitus other than T2DM and whether patients receive anti-hypertensive medication and/or aspirin were implemented in the model. However, these characteristics did not play a role, because the company assumed that no patients would have a history of angina or diabetes other than T2DM and no patients received anti-hypertensive medication and/or aspirin (see Section 5.2.3 for more details on baseline patient characteristics in the model).

If a patient experiences an event, the patient condition (or attribute as it is often called in DES terminology) for this event is updated. For example, if a patient is predicted to stop adjunct NB32 treatment before the first scheduled response assessment, a condition is used to record that the individual is no longer receiving adjunct treatment. Following updating of conditions, time to event (TTE) estimates are updated for any events affected by condition changes from the first event. For example, if an event changes BMI, times to obesity-related-disease events (for which BMI is an explanatory factor) are re-estimated.

In addition to the main modelling assumptions that are highlighted in Table 5.3 (retrieved from CS Table 53), the company's model assumed weight loss for orlistat patients at weeks 12 and 52 to be comparable to weight loss for NB32 patients at weeks 16 and 56. More specifically, the company assumed that the percentage weight loss (compared with baseline weight) for NB32 at weeks 16 and 56 can be combined with the mean difference between NB32 and orlistat (obtained from the ITC, see Sections 4.4 and 4.5 of his report for more details) to estimate the percentage weight loss (compared with baseline weight) for orlistat at weeks 12 and 52, respectively. This is similar for the proportion of responders at the week 16 and week 12 weight assessments (response criterion of ≥5% weight loss from baseline) for NB32 and orlistat, respectively. The company stated that the assumption of equivalent

weight loss at similar assessment times was also upheld within the ITC. No further justification is provided for this assumption.

Table 5.3: Main modelling assumptions utilised in the CS economic model

Assumption made	Rationale
Treatment discontinuation	
If a patient discontinues treatment with NB32 or orlistat, it is assumed that the patient is eligible to continue to receive non-pharmacological standard management (dependent on their sampled TTD).	Clinical expert consultation suggested that standard management would continue beyond cessation of adjunctive pharmacological therapy.
Weight regain	
Weight regain begins immediately after a patient discontinues all treatment (that is, adjunctive pharmacological treatment as well as standard management).	This assumption was made in the model built by Ara et al.[58] For patients who discontinue adjunctive therapy but continue to receive non-pharmacological standard management, weight regain was assumed to only commence when standard management was discontinued. Clinical expert opinion was sought to validate this assumption.
Weight is regained linearly over a 3-year period.	This assumption was made in the model built by Ara et al.[58]
The regained weight is reflective of the BMI expected as predicted by the natural history model for BMI over time.	BMI was assumed to revert to the natural history model predicted BMI given the intrinsic correlation known between age and BMI. This setting was included as a scenario analysis within the report by Ara et al.,[58] but was considered the most appropriate setting within the de novo model for incorporating BMI over time.
Obesity-related clinical events	
Within the model, it is possible for patients to experience a primary and secondary cardiovascular event (MI or stroke), as well as developing T2DM.	This assumption was made in the model built by Ara et al.[58] It is expected that the incremental clinical impact of further cardiovascular events would be negligible, as the proportion of patients who would experience more than two cardiovascular events in clinical practice is small.
Source: CS Table 53 Abbreviations: MI, myocardial infarction; T2DM, Type 2 diabetes mellitus; NB32, naltrexone 32mg plus bupropion.	

ERG comment: It is unclear to the ERG why a DES approach is preferred over for instance an individual-level state transition model. However, the ERG considered it reasonable to use the economic model by Ara et al.,[58] (comparing different pharmacological treatments for obesity) as a starting point for the current analysis. Based on their analyses, Ara et al.,[58] considered assumptions regarding weight regain to be key drivers of cost effectiveness. In this context, it should be noted that the company deviated from the assumption made by Ara et al.,[58] that patients would have regained weight to obtain their baseline BMI in three years and assumed instead that patients would have regained weight to obtain the predicted BMI in three years (using the natural history model predicting BMI over time by Ara et al.;[58] see Section 5.2.6 for more details). In response to clarification question B1a, the company responded that this was a 'logical' assumption for a simulated patient's BMI to be consistent with their characteristics.[9] However, also based on the responses to clarification question B1, it is illustrated that

this is not a conservative assumption for NB32 versus orlistat; the ICER vs. orlistat increased by £1,536 (Table 6 in the clarification response).[9] Moreover, the company did not provide justification for why their deviation from Ara et al.'s[58] assumption was 'logical' and more plausible than assuming weight regain to baseline BMI. After weight is regained to reach the baseline BMI, the BMI increases using the annual increase based on age (according to the correlation between age and BMI as reflected in the natural history model predicting BMI over time). Hence, to be consistent with Ara et al.,[58] and to be conservative, the ERG preferred to assume weight regain to the baseline BMI in its base-case. Furthermore, the linear weight regain over the time-span of three years was implemented incorrectly in the model where, in fact, the weight regain occurs instantaneously at the end of the three year period. The ERG incorporated adjustments in its base-case to reflect a linear weight regain over three year.

The ERG also questioned the (justification for the) assumption of equivalent weight loss at similar assessment times for NB32 and orlistat. The company's model assumed weight loss for orlistat patients at weeks 12 and 52 to be comparable to weight loss for NB32 patients at weeks 16 and 56. This was not justified besides stating this assumption was also upheld within the ITC (see Section 5.2.6 for more details).

The model only includes the possibility of two subsequent cardiovascular events (i.e. either two strokes, two MI's or one stroke and one MI). Implicitly assuming that the impact of the third cardiovascular event, on costs, quality of life and survival, is negligible. It can however be questioned whether having a stroke after having experienced two MIs is indeed unimportant. However, as the company argues in response to clarification question B9, this assumption is most likely conservative and hence considered reasonable by the ERG.

5.2.3 Population

NB32 is licensed as an adjunct to standard non-pharmacological management (i.e. reduced-calorie diet and increased physical activity) in adult patients who are obese (BMI ≥30kg/m2) or overweight (BMI ≥27kg/m2 and <30kg/m2) in the presence of one or more weight-related comorbidities (e.g., T2DM, dyslipidaemia, or controlled hypertension).[40] The company stated that the economic analysis aimed to reflect this patient group. Table 5.4 provides an overview of mean values for sampling baseline characteristics for individual patients in the model and used as explanatory factors for risks, costs or utility. According to the company these baseline patient characteristics were derived from a range of sources to best represent patients in UK clinical practice.

Table 5.4: Main modelling assumptions utilised in the CS economic model

Parameter	Mean value reported in CS	Mean value in economic model (calculated by ERG)	Justification	ERG value (if differently); based on section 4.2.2 (or stated if different)
Age	47 years	47 years	COR trial programme patient-level data	T2DM: 53.8 Non-T2DM: 44.7
Female	79.0%	76.7%		T2DM: 52.9% Non-T2DM: 86.7%
Height	Female: 1.64 m Male: 1.78 m	Female: 1.64 m Male: 1.78 m Total: 1.67		
Weight		Female: 90.3 kg		

Parameter	Mean value reported in CS	Mean value in economic model (calculated by ERG)	Justification	ERG value (if differently); based on section 4.2.2 (or stated if different)
	Derived from model	Male: 98.3 kg Total: 92.2 kg	Calculated by ERG based weight sampled in the model	
BMI	Derived from BMI trajectory model by Ara et al.[58] (see Section 5.3.4.3)	Female: 33.57 kg/m^2 Male: 31.05 kg/m^2 Total: 32.98 kg/m^2	Calculated by ERG based on height and weight sampled in the model	See Table 5.21 for BMI sampled in the ERG base-case
T2DM at baseline	33.2%	33.3%	Ara et al.[58]	
Insulin use for T2DM patients	33.3%	T2DM: 29.4% Total: 9.8%	Clinical opinion[64]	
Smoking status	Current: 7.0% Previous: 54.0% Never: 39.0%	Current: 5.7% Previous: 52.5% Never: 41.8%	Dare et al.[65]	Current: 10.6% Previous: 54.0% Never: 35.4%
Statin use	79.3%	80.4%	NB-CVOT study[29]	T2DM: 47.6% Non-T2DM: 10.4%
History of angina	0.0%	0.0%	Assumption – set to 0.0% as no data were identified for overweight/ obese patients	
History of diabetes other than T2DM	0.0%	0.0%		
receive anti-hypertensive medication	0.0%	0.0%	Assumption – set to 0.0% as it did cause counter-intuitive results	T2DM: 47.9% Non-T2DM: 15.0% (assuming antihypertensive medication in 77.3%[66])
receive aspirin	0.0%	0.0%		10.9% [58]
Source: CS Table 55 and economic model submitted by company Abbreviations: BMI, body mass index; DM, diabetes mellitus; T2DM, Type 2 diabetes mellitus.				

The clinical data used for NB32 and standard management during the first year in the economic evaluation are mainly retrieved from the four multicentre, randomised, double-blinded, placebo-controlled studies comprising the COR trial programme (COR-I, COR-II, COR-BMOD and COR-DM). In three of these studies (COR-I, COR-II, COR-BMOD) participants were adults with BMI 30–45kg/m2 or BMI 27–45kg/m2 and dyslipidaemia or controlled hypertension. In the fourth study (COR-DM), participants were adults with T2DM and BMI 27–45kg/m2. The company stated that, although no UK centres were included and the mean BMI of 36kg/m2 was slightly higher than usually seen in clinical trials, patient characteristics in the COR trial programme are a fair reflection of the typical patient group that would receive NB32 in UK NHS clinical practice.[64] This was based on clinical opinion (Professor

John Wilding, physician with extensive experience of treating overweightness and obesity in the NHS; JW).

Given the relatively limited follow-up period (56 weeks) of the trials in the COR trial programme and the necessity to project lifetime outcomes, the company used the NB-CVOT trial to estimate the outcomes (i.e. TTD) beyond the first year in the economic evaluation (748 patients receiving NB32 were followed beyond 52 weeks). The NB-CVOT study was a Phase IIIb, multicentre, randomised, double-blind, placebo-controlled trial to assess the occurrence of MACE in overweight or obese patients (randomising patients to receive treatment with NB32 or placebo).[29] Patients were eligible for inclusion if they were aged 45 (men) or 50 (women) years or older, had a BMI 27–50kg/m2 and a waist circumference of 88cm (women) or 102cm (men) or more. The company stated that for the NB-CVOT trial BMI inclusion criteria (BMI 27–50kg/m2 and a minimum waist circumference of 88cm (women) or 102cm (men)) were slightly different compared with the COR trial programme, patients in the NB-CVOT study were older than those in the COR trial programme (with inclusion restricted to men over 45 years and women over 50 years), and enrolment was restricted to patients with increased risk of cardiovascular outcomes.[31]

In addition to the COR trial programme and the NB-CVOT trial, the natural history model predicting BMI over time and risk equations developed by Ara et al.,[58] which predict lifetime BMI, risks for the development of key weight-related disease (i.e. stroke, MI and T2DM) and death, were used in the economic model. This was based on adult patients from the GPRD (General Practice Research Database; accessed in January 2011) who had three or more BMI readings of over 27kg/m2 (see Section 5.2.6 for more details).

ERG comment: The population aimed to reflect the scope.[7] However, patient characteristics in the model were sampled from estimates that were based on a variety of sources. It is questionable whether this is reflective of UK clinical practice. The ERG agrees with using the COR trial programme patient-level data to inform baseline age, gender and height in the model. This follows from a) that the effectiveness estimates are derived from this population and b) that the company stated, based on clinical opinion (JW), that patient characteristics in the COR trial programme are a fair reflection of the typical patient group that would receive NB32 in UK NHS clinical practice.[64] However, the other baseline characteristics can be questioned.

The ERG compared the baseline BMI sampled in the model with the baseline BMI in the COR trial programme (Table 5.5). This comparison indicates that baseline BMI is vastly underestimated in the economic model, compared to the COR trial programme and as such also compared to UK clinical practice (as clinical opinion indicated that patient characteristics in the COR trial programme are a fair reflection of the typical patient group that would receive NB32 in UK NHS clinical practice). This is also reflected in the average baseline weight of 92kg sampled in the model, while the averages ranged between 99kg and 105kg in the COR trial programme (see Section 4.2.2 for more details). Given that BMI is included as a predictive factor for utility, T2DM, cardiovascular events and death (see sections 5.2.6 and 5.2.8 for more details), the utility values and the time to these events in the model are overestimated, likely inducing bias in favour of NB32.

Other baseline characteristics are also potentially underestimated:

- Proportion current smoker of 7% (sampled 6%) while the averages ranged between 9% and 11% in the COR trial programme (excluding COR-BMOD as this trial included only non-smokers; see Section 4.2.2 for more details).

- Proportion receiving anti-hypertensive medication of 0% while the averages of hypertensive patients ranged between 15% and 63% in the COR trial programme (see Section 4.2.2 for more details). Moreover, after reviewing the time to obesity-related events, it is unclear why the company indicated that setting this to >0% would lead to counter-intuitive results.
- Proportion of patients with history of angina and/or diabetes other than T2DM of 0%. Although there were hospitalisations for unstable angina (see Section 4.2.7), the company stated that there were no data to inform these proportions.
- Proportion receiving aspirin of 0% while based on Ara et al.[58] this can be calculated to be >10%. Moreover, after reviewing the time to obesity-related events, it is unclear why the company indicated that setting this to >0% would lead to counter-intuitive results.

In addition to this, the GPRD population from Ara et al.[58] had three or more BMI readings of over 27kg/m2, but this population did not consider whether patients had one or more weight-related comorbidities while NB32 is licensed for patients with a BMI between 27-30kg/m2 only in the presence of one or more weight-related comorbidities. Hence, it is the ERG's view that both the baseline patient characteristics and the risk equations developed by Ara et al.[58] to predict lifetime BMI and risk for the development of key weight-related diseases are based on a less severe population than the licensed indication for NB32.

In contrast to the above, the proportion of patients with diabetes might have been overestimated. The value of 33.3% was obtained from Ara et al., but this was not validated against the population in the scope i.e. those with BMI ≥30kg/m2) or overweight (BMI ≥27kg/m2 and <30kg/m2) in the presence of one or more weight-related comorbidities (e.g., T2DM, dyslipidaemia, or controlled hypertension). According to Health Survey England data for 2013, the percentage is between 14 and 15% depending on sex.[67]

Also, the proportion of patients receiving statins might have been overestimated as this is 8% to 13% in the COR-I, COR-II and COR-BMOD trials and 46% to 49% for diabetic patients in the COR-DM trial (see Section 4.2.2) while it was 79% (sampled 80%) in the economic model independent of T2DM status. Related to this, correlations between covariates were not incorporated in the sampling of the patient characteristics, leading to counter-intuitive patient profiles. For instance, based on the patient characteristics table of the COR-I, COR-II, COR-BMOD and COR-DM trials (in Section 4.2.2), it becomes clear that the patients without T2DM (COR-I, COR-II and COR-BMOD trials) have different patient characteristics (e.g. regarding age, sex, hypertension status and statin use) than patients with T2DM (COR-DM trial). This is neglected in the sampling of the patient population. To address these issues, the ERG adjusted the baseline characteristics used in the model (Table 5.4). This included calibrating the natural history model to predict BMI over time (see Section 5.3 for more details).

The company assumed no patients had a history of angina and/or diabetes other than T2DM. This assumption was made as no data were identified on these characteristics for overweight/obese patients. The ERG agrees with this statement and would therefore argue that it can be questioned whether the results of the economic analyses are representative for patients with a history of angina and/or diabetes other than T2DM.

Table 5.5: Distribution of BMI (patients sampled in the model and across the COR trial programme)

Obesity class	Model			COR-I		COR-II		COR-BMOD		COR-DM		NB-CVOT	
	Overall	Female	Male	Placebo	NB32	Placebo	NB32	Placebo	NB32	Placebo	NB32	Placebo	NB32
BMI<30kg/m2	8.0%	3.3%	23.6%	0.9%	3.1%[a]	2.8%	2.5%	0.5%	1.4%	6.5%	5.4%	7.0%	6.7%
BMI ≥30 and ≤35 kg/m2	74.6%	74.2%	76.0%	37.3%	38.4%	37.6%	39.8%	31.7%	35.0%	28.8%	33.1%	31.6%	31.3%
BMI ≥35 and <40 kg/m2	17.1%	22.2%	0.4%	39.4%	35.0%	38.6%	31.6%	39.1%	38.9%	37.6%	32.8%	31.1%[a]	33.2%[a]
BMI ≥ 40 kg/m2	0.3%	0.4%	0.0%	22.4%	23.5%	21.0%	26.2%	28.7%	24.7%	27.1%	28.7%	30.3%	28.8%

Source: Economic model submitted by the company and response to clarification question A17

[a]Original value in response to clarification question A17 contained incorrect proportions, this is corrected by the ERG.

5.2.4 Interventions and comparators

In line with the final scope and licensed indications, the company considered the following comparators for NB32 as an adjunct to standard management:

- orlistat as an adjunct to standard management and;
- standard management alone

NB32 is implemented as per its EMA Summary of Product Characteristics (SmPC) posology and method of administration, incorporating a four week escalation period, after which the maximum recommended daily dose of 32mg naltrexone hydrochloride and 360mg bupropion hydrochloride is assumed.[40] Orlistat is similarly implemented as per its EMA SmPC posology and method of administration, a 360mg daily dose.[39]

The company specified standard management as implemented in the analysis to reflect the non-pharmaceutical dietary and lifestyle management treatment received in UK NHS practice (see Section 5.2.9 for more details). The company stated based on clinical opinion (JW) that although standard management varies by geography, the non-pharmaceutical treatment administered in the COR-I and COR-II studies is a good reflection of the treatment patients are likely to receive in NHS England.[64] According to the NB32 license, standard management includes a reduced-calorie diet and increased physical activity.[40]

Stopping rules for both NB32 and orlistat are implemented in the model, as per their license terms:[39]

- NB32: patients who fail to meet the response criterion of ≥5% weight loss from baseline after 16 weeks after treatment initiation (12 weeks post-escalation period) discontinue pharmacological treatment.
- orlistat: patients who fail to meet the response criterion of ≥5% weight loss from baseline after 12 weeks after treatment initiation, discontinue pharmacological treatment.

Based on clinical opinion (JW),[64] the same stopping rule was applied 56 and 52 weeks after treatment initiation for NB32 and orlistat, respectively. It should be noted that these stopping rules only apply to pharmacological treatment (not necessarily to standard management that is provided in addition to NB32/orlistat), see Section 5.2.6 for more details regarding TTD.

ERG comment: The ERG considered whether, given that it is not required according to the license terms,[39, 40] the stopping rule at the secondary assessment, i.e. at 56 and 52 weeks after treatment initiation for NB32 and orlistat, would be reflective of clinical practice. The ERG found that in NICE clinical guideline 189 regarding obesity,[5] it is recommended (Section 1.9.9) that there will be a discussion regarding drug treatment longer than 12 months after discussing potential benefits and limitations. It should however be noted that this recommendation does not consider an objective response criterion such as the ≥5% weight loss from baseline used in the model.

One major limitation of the model is the inability to incorporate re-treatment, behaviour modification treatment (e.g. a weight loss programme) and or bariatric surgery (for which patients become eligible over time once their BMI is/increases to >40kg/m2[68] in the model). The company stated (in response to clarification question B3) that re-treatment is clinically plausible. However, the company did not incorporate this justified by a stated lack of data to inform re-treatment in the model.

5.2.5 Perspective, time horizon and discounting

The analysis was conducted from the perspective of the payer, i.e. the NHS England and Wales, over a lifetime horizon. Costs and outcomes were discounted by 3.5%.

ERG comment: This is in line with the NICE reference case.

5.2.6 Treatment effectiveness and extrapolation

i) Treatment effectiveness and extrapolation overview

In the CS,[1] clinical parameters and variables are reported as falling into the following four categories:

- Baseline patient characteristics,
- Treatment duration,
- Treatment effectiveness,
- Epidemiological models of natural history.

In this report, the baseline patient characteristics were presented and discussed in Section 5.2.3. Time to treatment discontinuation is discussed in Section 5.2.6 ii). Treatment effectiveness is discussed under the headings iii) Proportion of patients with weight loss ≥ 5% and iv) Mean change in body weight. Finally, obesity-related events and epidemiological BMI models are discussed in Section v) Risk of obesity-related events and natural history of BMI.

Treatment effectiveness estimates (i.e. time to treatment discontinuation data, proportion of responders, and mean change in body weight) were derived from the COR trial programme, including the COR-I, COR-II, COR-BMOD and COR-DM trials. All the analyses were based on the company's mITT analysis, which reflects only those patients who have a post-baseline measurement whilst on the study drug.

ERG comment:

The ERG questions the company's approach 1) using the mITT analysis for estimating the proportion of responders and mean change in weight and; 2) pooling across all COR studies for estimating the time to treatment discontinuation, proportion of responders and mean change in weight.

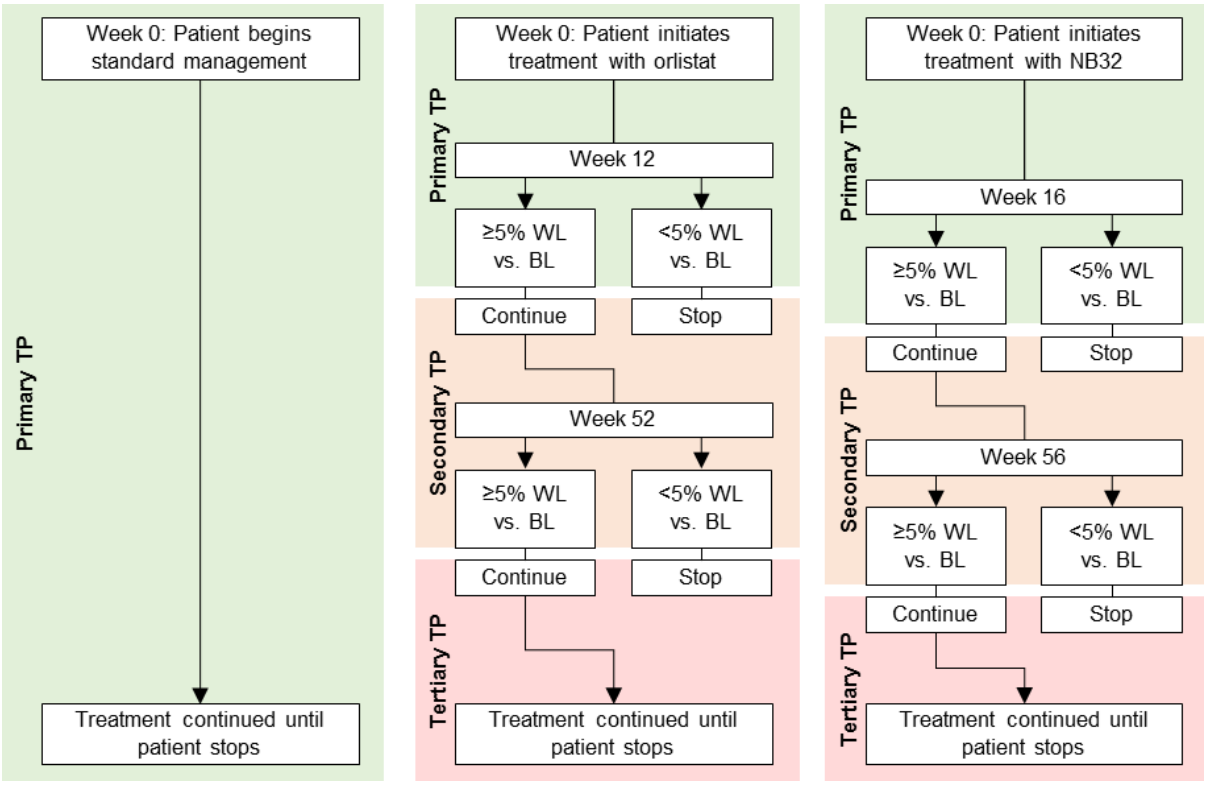
The ERG considers that the use of the ITT population would have been both more appropriate and more conservative. The mITT analysis includes only those patients who have a baseline and at least one post-baseline measurement whilst on the study drug. Patients who discontinued without providing follow-up weight assessments were excluded. Reasons for discontinuation could relate to efficacy or safety (i.e. AEs) of the drug. Using the true ITT data, NB32 would achieve a smaller mean percentage of weight loss and smaller proportion of responders compared to the mITT data. This is discussed in more detail in Section 4.2. Following the ERG's request for scenario analyses using data on clinical effectiveness and treatment discontinuation derived from the ITT population (Question B6), the company refused to carry out these analyses, stating that these were *"irrelevant to de novo model, due to the nature in which weight loss outcomes are derived"*.[9] Whilst the company justified this by clarifying that the safety population, not the mITT population, was used to estimate time to treatment discontinuation (TTD) up to week 16, no further justification was provided for not presenting the scenario analyses using ITT estimates for proportion of responders and mean change in weight. The issue of using this population and the bias that it introduces are discussed further in Sections 5.2.6 iii) and iv).

It is the ERG's view that it was inappropriate to pool from all COR studies, including COR-BMOD and COR-II. In the CS, it is stated that, according to the company's criteria, the COR-BMOD study considered 'intensive' behaviour modification. This was based on *"the number of follow-up appointments with a medical/dietary professional; detail and severity regarding the prescription of dietary recommendations; and the level of physical activity participants were encouraged to follow"*.[1] In the response to clarification question B8[9], the company states that based on clinical expert opinion, the effects of intense behavioural modification and pharmacological treatment would be expected to be additive. This might suggest that the difference between pharmacologic treatments would remain the same irrespective of the intensity of non-pharmacological treatment. However, effectiveness estimates derived from COR-BMOD, where NB32 was administered in combination with intensive behavioural modification were substantially different when compared to effectiveness estimates derived from the studies in which NB32 was administered together with standard management only. Pooling clinical effectiveness data from all COR trials, including the COR-BMOD study, is therefore inappropriate. The ERG notes that if effectiveness estimates included intense behavioural modification, then this should also be reflected in the cost. In the absence of cost estimates, the ERG was unable to perform analysis including intense behavioural modification. Likewise, the ERG considers the use of COR-II for the derivation of treatment effectiveness beyond 28 weeks as inappropriate because NB32 participants with <5% weight loss at visits between weeks 28 and 44 were re-randomised. The ERG therefore considers that NB32 treatment effectiveness estimates, assuming no concomitant behaviour modification (e.g. weight loss programme), should only be derived from the COR-I and COR-DM trials. The issue of pooling from all COR studies and the bias that it introduces are discussed further in Sections 5.2.6 ii)-iv).

ii) Time to treatment discontinuation

Time to treatment discontinuation was estimated separately for patients receiving standard management, NB32 and orlistat. For patients receiving standard management (alone or in combination with adjunctive pharmacological therapy), treatment is given from week 0 until the patient stops treatment. For patients receiving pharmacological therapy, treatment duration is considered in three phases: phase 1 includes the time to primary assessment (conducted at week 16 for NB32 and at week 12 for orlistat); phase 2 is the time from primary to secondary assessment (which is conducted at week 56 for NB32 and at week 52 for orlistat); and phase 3 covers the time after the secondary assessment. It is stated within the CS that *"patients must cease to receive adjunctive therapy ahead of discontinuing standard management, after which they may either immediately discontinue standard management or continue to receive standard management alone"*.[1] Kaplan–Meier (KM) data and the proportion of responders are used to inform the duration of adjunctive therapy and standard management within the model. The expected pathway of care is illustrated in Figure 26 of the CS,[1] printed here in Figure 5.2.

Figure 5.2: Expected pathway of care across all treatment arms



Key: BL, baseline; NB32, naltrexone 32mg plus bupropion; TP, treatment phase; WL, weight loss.

Phase 1 (from treatment initiation to primary assessment):

For both NB32 and orlistat patients, treatment duration in phase 1 was based on KM estimates from NB32 treatment discontinuation data in the COR trial programme (illustrated in Figures 5.3 and 5.4). These data were also used for orlistat because *"there were no comparable duration of treatment data available to inform discontinuation ahead of primary assessment…"*.[1] However, because phase 1 was shorter for orlistat than for NB32 (12 weeks instead of 16 weeks), the KM data for NB32 patients were linearly scaled to fit the 12-week period to primary response assessment for orlistat. For NB32 patients, 67.2% continued treatment until 16 weeks. As a result of the linear scaling, this same proportion was also used for orlistat at 12 weeks.

Figure 5.3: NB32 adjunct therapy discontinuation from treatment initiation to 16 weeks (pooled COR trial programme data, all NB32 patients)

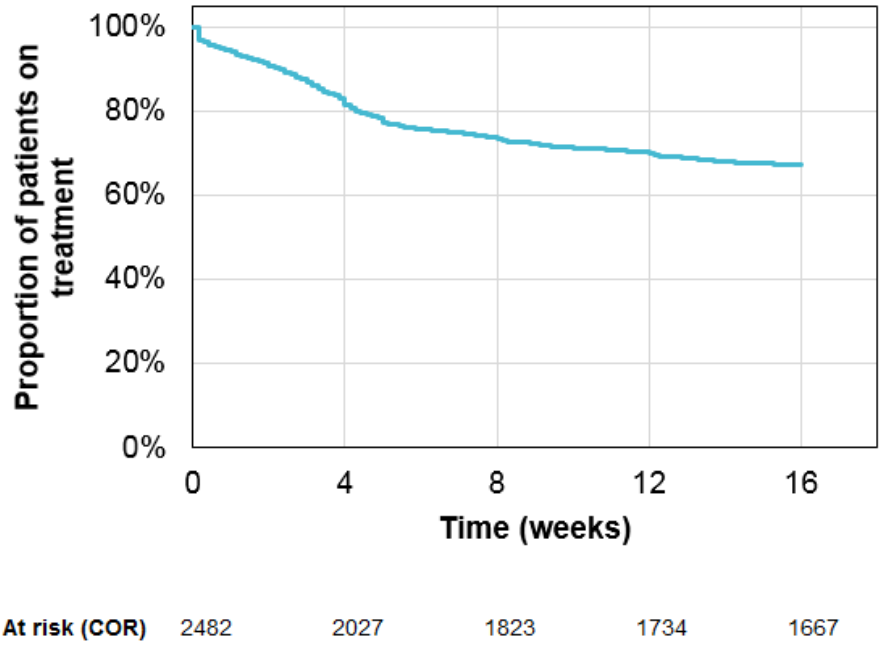
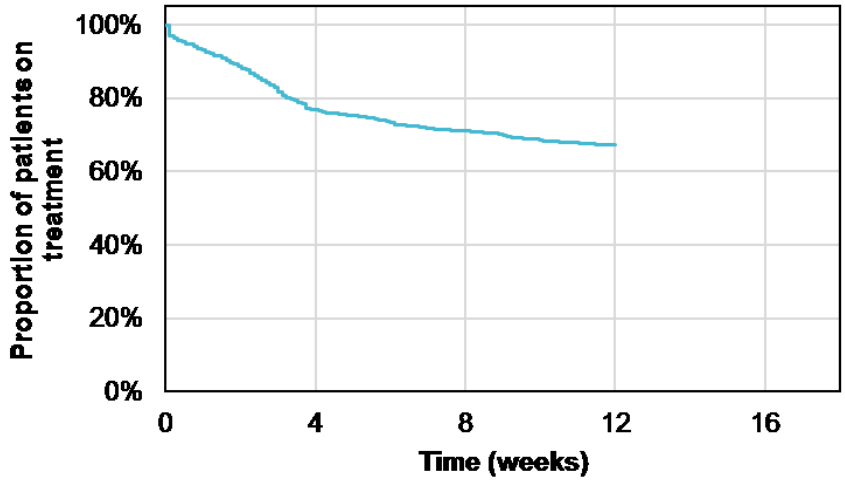


Figure 5.4: Orlistat adjunct therapy discontinuation from treatment initiation to 12 weeks (from pooled COR trial programme data, all NB32 patients)



Phase 2 (from primary assessment to secondary assessment):

For both NB32 and orlistat patients, treatment duration in phase 2 was based on KM estimates from NB32 treatment discontinuation data in the COR trial programme, with only those patients included in the analysis that had achieved response at their primary assessment date (i.e. a weight loss of at least 5% compared with baseline). Because treatment discontinuation data were not available for orlistat, the same NB32 treatment discontinuation KM data were used for orlistat, but shifted by four weeks to match the shifted time from primary to secondary assessment (12 to 52 weeks instead of 16 to 56 weeks). For NB32 patients, 86.1% of responding patients at week 16 continued treatment until 56 weeks. For orlistat patients, the same proportion of responding patients continued treatment until 52 weeks. This is illustrated in Figures 5.5 and 5.6 below.

Figure 5.5: NB32 adjunct therapy discontinuation from 16 to 56 weeks (from pooled COR trial programme data; NB32 16-week responders)

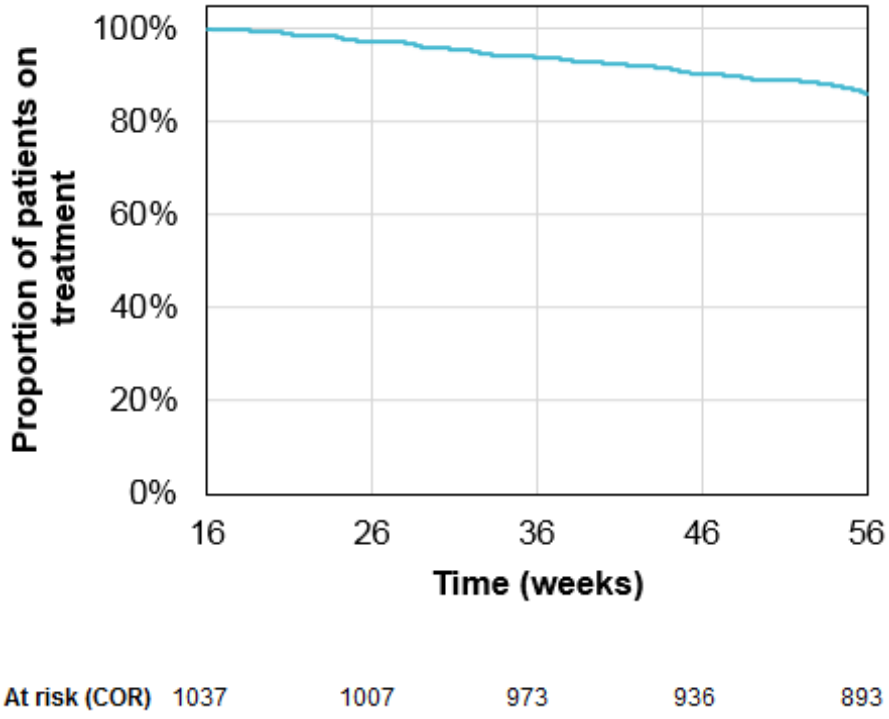
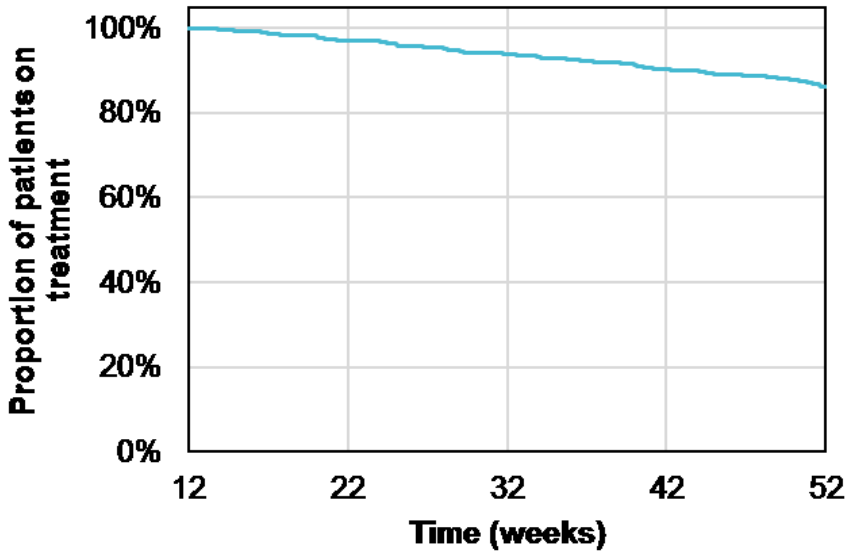


Figure 5.6: Orlistat adjunct therapy discontinuation from 12 to 52 weeks (from pooled COR trial programme data; NB32 16-week responders)



Phase 3 (from secondary assessment onwards):

For both NB32 and orlistat patients, treatment duration in phase 3 was based on KM estimates from NB32 treatment discontinuation data in the NB-CVOT study for the time period from 56 weeks to 158 weeks (end of study period). All patients were assumed to have discontinued after treatment duration data were unavailable (see Figure 5.7 and 5.8 below).

Figure 5.7: NB32 adjunct therapy discontinuation from 56 weeks (from NB-CVOT study data; NB32 56-week responders)

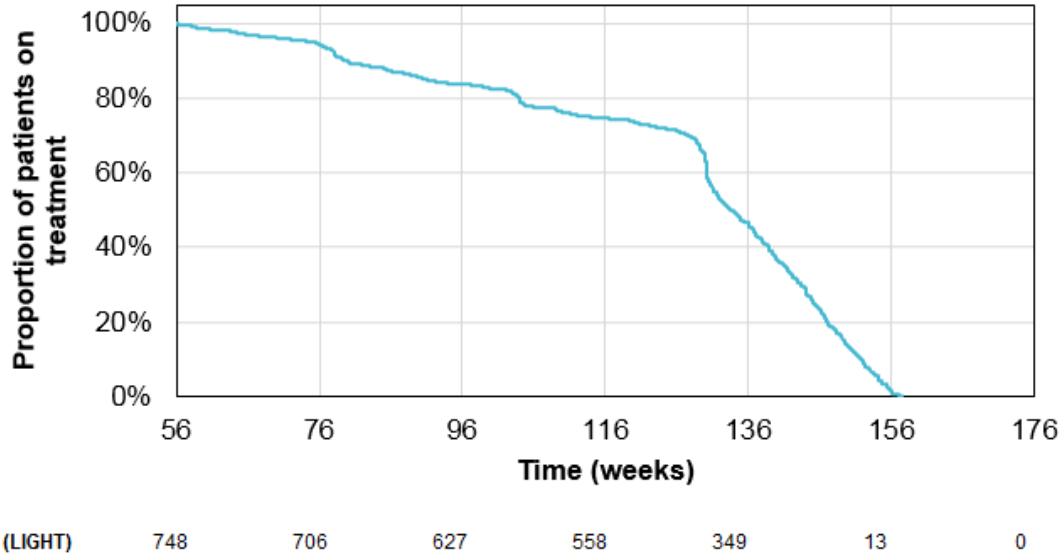
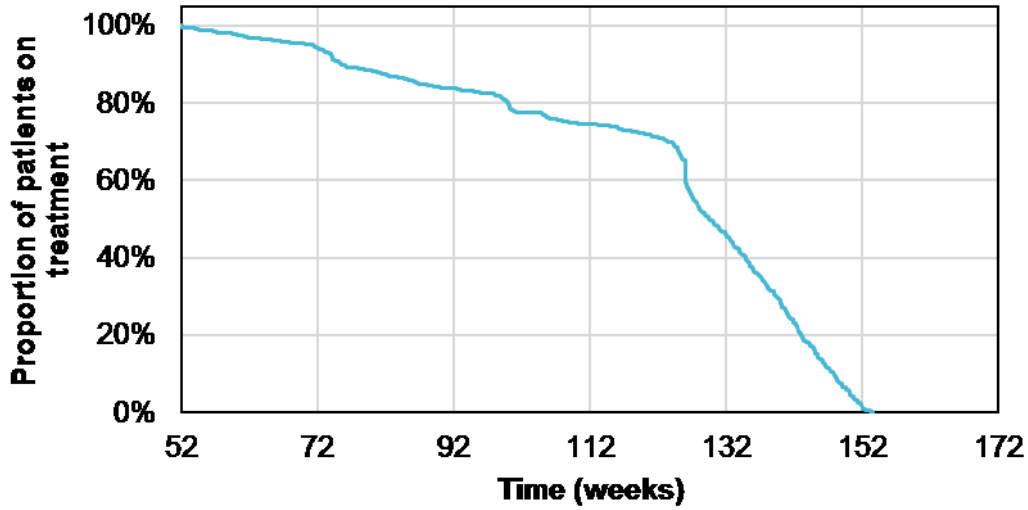


Figure 5.8: Orlistat adjunct therapy discontinuation from 56 weeks (from NB-CVOT study data; orlistat 52-week responders)

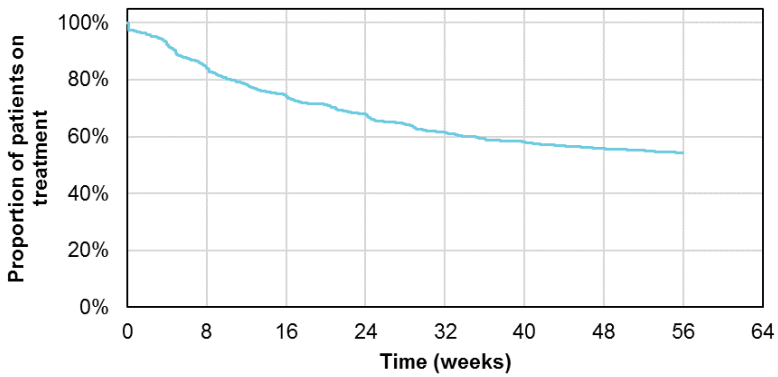


Treatment duration estimation for standard management:

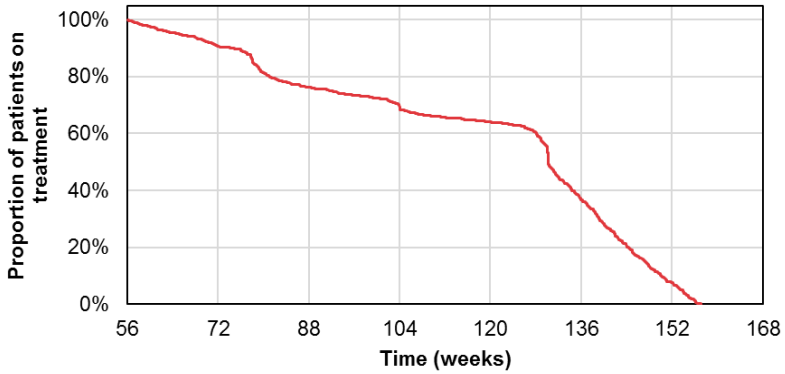
Patients receiving standard management are not subject to the same response-based treatment stopping rules as those receiving adjunct pharmacological treatment. Therefore all patient-level data from the COR trial programme could be used to inform TTD estimates in the first 56 weeks after treatment initiation. Treatment duration for standard management was then estimated using the available data from the COR trial programme up to 56 weeks and then joining the KM data from NB-CVOT to KM data from the COR trial programme by scaling the curve according to the proportion of patients who were still receiving standard management treatment at week 56 (see Figure 5.9 below).

Figure 5.9: Derivation of duration of standard management treatment

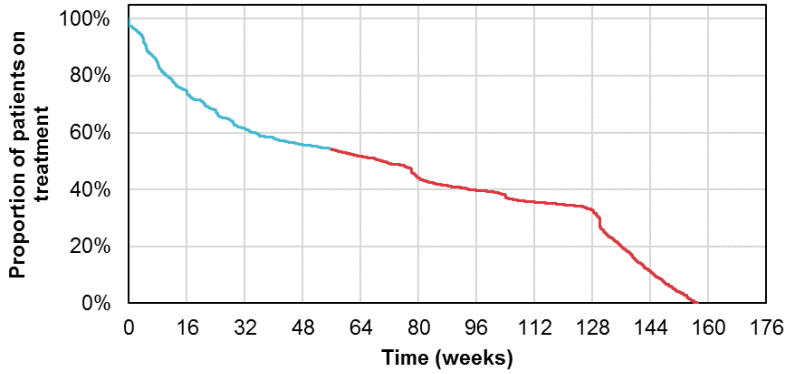
Kaplan–Meier data from the COR trial programme (Week 0 to Week 56)



Kaplan–Meier data from the NB-CVOT study (Week 56 onwards)



Combined Kaplan–Meier data from both studies (Week 0 onwards)



ERG comment: The ERG considers the use of the safety population for TTD as reasonable but believes that the TTD is underestimated in the model, in particular for orlistat.

Based on the CS,[1] it was unclear which population was used to estimate TTD. The company clarified in their response to Question A19 of the clarification letter[9] that the safety population was used to estimate TTD up to week 16. In the CS (Section 4.4), it is stated that the safety population *"included all randomised patients who were administered at least one tablet of study treatment and had at least one investigator contact/assessment at any time after the start of study treatment, regardless of whether they discontinue the study"*.[1] The ERG wishes to highlight that the ITT definition commonly used in orlistat trials is closer to this definition of the safety population in COR trials than to the company's mITT population used for the COR trial programme.[42-44] The company argued in their response to clarification question A19.b that the KM function of ITT and safety populations would be the same except for the number of patients at risk at time 0. This difference would stem from untreated patients who would automatically be censored at time 0. According to the company, this would make the two KM functions equivalent. The ERG was satisfied that this was reasonable.

It is the ERG's view that TTD should not have been pooled from the four studies in the COR trial programme. The ERG considers it to be plausible that treatment duration is different in patients who also receive intense behavioural modification. Furthermore, TTD may be different in patients with or without T2DM. The ERG therefore considers that modelling TTD separately for both subgroups (T2DM and non-T2DM) may have been more appropriate.

The patient output from the company's model run revealed a mean TTD of 13.32 months, 12.29 months and 17.16 months for NB32, orlistat and for SM respectively. The ERG thinks that these may be under-estimates because:

(1) TTD estimates for the period after the one year assessment were derived from the NB-CVOT study in which patients had characteristics associated with an increased risk of CV outcomes, potentially leading to a shorter TTD. This is acknowledged by the company in the CS, in which it is stated that TTD is likely to be under-estimated by these data, given the age and comorbidity profile of NB-CVOT study patients[1].
(2) The end of the NB-CVOT study was used as the maximum TTD, whether patients in that study had discontinued or not.
(3) The company claims that the most reasonable and conservative assumption was to assume that TTD for orlistat would follow a similar trajectory to NB32, given that patient-level data for orlistat were unavailable. However, the ERG found publications reporting TTD for orlistat, which reveal that orlistat TTD was longer than the 12.29 months estimated by the model, with many studies reporting that the proportion of patients still receiving orlistat at 12 months was >50%.[42-44] However, these TTD estimates were not conditional on response to treatment (primary and secondary response assessments) and therefore have to be interpreted with caution, but reported response rates in two of these studies suggest that a significant proportion would still have continued treatment based on their response (45.7% response rate as measured by patients achieving >5% weight loss in Bakris et al.[42], 55.6% in Broom et al.[44]; Berne et al.[43] did not report response rates with the same level of weight loss). It is the ERG's view that the company should have validated their assumption for orlistat with these data. Furthermore, the company claimed that TTD may be shorter with orlistat than with NB32, given the known toxicity profile and association with treatment discontinuation in Question B2 of the clarification response.[9] This is, however, not supported by any evidence. It is the ERG's view

that TTD for NB32 and orlistat may be under-estimated. The ERG wishes to highlight that the under-estimation of TTD leads to an under-estimation of costs.

(4) For the derivation of the orlistat TTD, the KM estimates for NB32 TTD for the first 16 weeks were linearly scaled to fit the first 12 weeks of orlistat treatment. This was justified by the different time to primary assessment, and the fact that for NB32, the first four weeks include a titration period. The ERG believes that this linear scaling may further under-estimate orlistat TTD, resulting in worse effectiveness (patients will stop losing weight and start weight regain sooner), but also in decreased costs associated with orlistat and the effect of this is therefore ambiguous. The ERG therefore removed the linear scaling in its base-case analysis. The ERG furthermore considers there to be considerable uncertainty surrounding the TTD of orlistat estimation.

iii) Proportion of patients with weight loss ≥ 5%

The proportion of patients with weight loss ≥ 5% at primary response assessment (conditional on being on treatment) was obtained for NB32 by dividing the proportion of patients (50.8%) still on treatment by the total proportion of patients (65.2%) that had achieved a ≥ 5% weight loss at primary response assessment in the COR trial programme (COR-I, COR-II, COR-BMOD, COR-DM). The proportion of patients continuing treatment after this assessment, was thus estimated to be 75.7% of those still on treatment.

For orlistat, the proportion of patients with weight loss ≥ 5% at primary response assessment was not available. The company therefore used the relative effectiveness estimate for proportion of responders at secondary response assessment at one year (which was available as an odds ratio derived from the ITC) to obtain the proportion of responders at primary assessment. This yielded a proportion of 77.9% for the T2DM and 70.5% for non-T2DM groups, respectively.

At secondary response assessment, mean change in body weight estimated in the model determines the proportion of responders and non-responders.

ERG comment:

The ERG notes that there was a discrepancy between the mean OR for the proportion of responders of orlistat compared with NB32 used in the model and the one reported in the CS on page 19 and in a forest plot shown in Figure 19 of the CS.[1] In the CS, a mean OR of 1.09 is reported, whilst the model uses a mean OR of 1.13, which is based on the coda sample. It is important to note that both of these values would mean that a greater proportion of patients would achieve weight loss of ≥5% with orlistat compared to NB32 at the one year assessment. The company, however, notes that that difference based on the mean OR of 1.09 was not statistically significant. It was unclear whether a mistake was made in the report or within the model (the coda sample used) and the ERG was therefore unable to address the discrepancy. The ERG, however, notes that, if the mistake was in the model, then this would have likely caused a slight upwards bias to the ICER comparing NB32 with orlistat. Furthermore, this discrepancy is addressed in the ERG's base-case analysis where the ITT data and therefore a newly calculated OR is used.

The ERG's concerns about the derivation of proportion of responders for NB32 and comparators are presented in the following paragraphs for NB32 (1) and comparators (2):

(1) As was stated above in Section 5.2.6 i), it was inappropriate to pool the proportion of responders to NB32 treatment from all COR studies, including BMOD. By doing so, the proportion of responders to NB32 is over-estimated. This is supported by response rates for treatment with NB32 versus placebo as

presented in Table 4.8. As a result, it is the ERG's view that response rates to NB32 are likely to be over-estimated as a consequence of the pooling method.

Furthermore, the use of mITT data for the derivation of response rates would bias the estimates in favour of NB32. This is shown in Table 4.10 of the clinical effectiveness section, which shows that a smaller proportion of patients achieve a response when the ITT population is used, compared with the mITT population. The company was asked to provide an analysis using ITT populations but failed to do so.

(2) The ERG considers that the application of the base-case odds ratio derived from the ITC is also inappropriate because this was derived from all four COR studies, including the COR-BMOD and COR-II studies. The more appropriate estimation of both NB32 and orlistat rate of responders would be to use the rate of responders as pooled from COR-I and COR-DM and then apply the odds ratio derived from sensitivity analysis 3 in the ITC, which excludes the studies in which pharmacological treatment is combined with more intensive behavioural modification. The ERG also wishes to highlight that the estimation of the orlistat response rate at primary assessment was made based on the assumption that the one year odds ratio between orlistat and NB32 equally applies to the 12/16 week setting. The ERG was satisfied that, in the absence of other data, this was a reasonable assumption.

iv) Mean change in body weight

For NB32, mean change in body weight was estimated separately for responders and non-responders at the primary response assessment (16 weeks) and derived from the COR trial programme (COR-I, COR-II, COR-BMOD, COR-DM).

For orlistat, mean change in body weight compared with NB32 was derived from the ITC, assuming that weight loss at 16 weeks in NB32 patients was comparable with weight loss at 12 weeks in orlistat patients. This assumption was justified in the CS by the lack of a four week titration period for patients treated with orlistat. Moreover, due to lack of weight loss data for orlistat at 12 weeks and due to it not being possible to stratify weight loss by response status in the ITC, the relative weight loss (as in the mean difference in weight loss) of orlistat compared with NB32 at the one year assessment was used to estimate weight loss associated with orlistat treatment at 12 weeks (primary response assessment) for both responders and non-responders. For standard management patients, weight loss estimates were derived from the COR trial programme patient-level data and not stratified by response. For both orlistat and standard management patients, weight loss estimates were stratified by T2DM status, but for NB32 this was not done.

The average weight loss data used in the model are summarised in Table 5.6.

Table 5.6: Average weight loss at primary response assessment

Treatment	Outcome	Value	Source
NB32	Primary Week 16 assessment: Responders	9.4%	COR trial programme data
	Primary Week 16 assessment: Non-responders	1.9%	
ORL	Primary Week 12 assessment: Responders (all patients)	8.6%[a]	ITC[a]
	Primary Week 12 assessment: Responders (T2DM patients)	9.2%	ITC
	Primary week 12 assessment: Responders (non-T2DM patients)	8.3%	
	Primary Week 12 assessment: Non-responders (all patients)	1.1%[a]	ITC[a]

Treatment	Outcome	Value	Source
	Primary Week 12 assessment: Non-responders (T2DM patients)	1.7%	ITC
	Primary Week 12 assessment: Non-responders (non-T2DM patients)	0.8%	
SM	Week 12: All patients	2.3%	COR trial programme data
	Week 16: All patients	2.7%	
ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SE, standard error; SM, standard management; T2DM, Type 2 diabetes mellitus. Notes: [a], The derived proportion shown here is an estimate based upon the proportion of T2DM patients at baseline.			

At secondary response assessment, weight loss for NB32 patients was calculated as 11.7% for those who responded at primary response assessment and 8.8% for all patients combined. For orlistat, the 11.7% weight loss for NB32 patients was used together with the mean difference in weight loss of orlistat compared to NB32 derived from the ITC to estimate weight loss for orlistat responders at 52 weeks (weight loss for NB32 patients at week 56 was again assumed to be comparable to that of orlistat patients at week 52 given the lack of a four week titration period for orlistat). For standard management, weight loss was estimated based on the weight loss calculated for all NB32 patients in the COR trial programme regardless of response status (of 8.8%, as stated above) and the mean difference in weight loss based on the ITC, stratified by T2DM status. The average weight loss figures for the secondary response assessment are presented in Table 5.7.

Table 5.7: Average weight loss at secondary response assessment

Treatment	Outcome	Value	Source
NB32	Secondary Week 56 assessment	11.7%	COR trial programme data
ORL	Secondary Week 52 assessment (all patients)	10.9%	ITC[a]
	Secondary Week 52 assessment (T2DM patients)	11.5%	ITC
	Secondary Week 52 assessment (non-T2DM patients)	10.6%	
SM	Week 52/56 (all patients)	4.5%	ITC[a]
	Week 52/56 (T2DM patients)	5.6%	ITC
	Week 52/56 (non-T2DM patients)	3.9%	
ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SE, standard error; SM, standard management; T2DM, Type 2 diabetes mellitus. Notes: [a], The derived proportion shown here is an estimate based upon the proportion of T2DM patients at baseline.			

ERG comment: As was stated above in Sections 5.2.6 i) and iii), it was also inappropriate to pool the mean change in weight from all COR studies, including BMOD. By doing so, the proportion of responders to NB32 is over-estimated (see Table 4.8). Similarly, mean weight loss figures in the COR-BMOD study are larger in both the NB32 and placebo arms than in the other studies. As a result, it is the ERG's view that mean change in weight for patients treated with NB32 is likely to be over-estimated as a consequence of the pooling method (Table 4.8).

The use of mITT data for the derivation of mean change in weight introduces further bias in favour of NB32. This is shown in Table 4.9 of the clinical effectiveness section, which shows that a smaller mean change in weight is achieved when the ITT population is used, compared with the mITT population. The company was asked to provide an analysis using ITT populations but failed to do so, without providing adequate justification.

The ERG calls into question the assumption of weeks 12 and 52 on treatment with orlistat being comparable to weeks 16 and 56 for patients treated with NB32. The justification provided by the company was that the first four weeks of treatment with NB32 were a titration period. However, patients do lose weight even during this titration period, as shown in Figure 7 in the CS[1], where patients lose most weight in the first four weeks, followed by a phase where weight loss slows down. The ERG therefore considers this assumption as inappropriately justified. However, based on the same figure, patients may not lose significantly more weight in the last four weeks of a one year treatment period. Therefore the assumption of equivalence appears more valid for the one year assessment than for the primary response assessment.

The ERG wishes to highlight a further limitation in the company's analysis of weight loss for patients treated with orlistat at 12 weeks. This was derived by using the mean difference in weight loss of orlistat compared with NB32 as derived from the ITC at the one year assessment. The mean difference is an absolute measure, which would presumably vary according to the magnitude of weight loss achieved. With absolute weight loss being smaller at the primary than at the secondary response assessment at one year, applying the absolute mean difference derived from one year to NB32 weight loss data will result in an under-estimation of weight loss for patients treated with orlistat. The ERG therefore adjusted this in its base-case.

v) Risk of obesity-related events and natural history of BMI

In the CS, the risk of occurrence of obesity-related events is modelled conditional on changes in BMI, whereby BMI levels are assumed to change with age, based on a study by Ara et al.[58]. This is achieved in two parts:

1. Through the use of risk equations to estimate the time to stroke, MI, onset of T2DM and death.
2. Through a natural history model of BMI over time to estimate patients' BMI throughout their lifetime.

In the choice of the risk equations and BMI natural history model, the company heavily relied on the previously published HTA report by Ara et al.[58] The company states that the study by Ara et al.[58] identified limitations with existing studies of the relationship between the development of cardiovascular disease and weight. Those studies comprised cross-sectional studies identifying correlations between major clinical events and BMI, studies that categorised BMI and were therefore unable to capture changes within categories and other existing studies primarily conducted outside the UK. Ara et al.[58] therefore used large-scale GPRD data to estimate the risk of major cardiovascular events occurring at specific levels of BMI and age, controlling for confounding factors.

To establish both the risk equations for major clinical events and the natural history model of BMI, Ara et al.[58] drew 100,000 patients with three or more BMI readings over 27kg/m^2 from the GPRD database. For 1. the risk equations for obesity-related events, occurrence of all-cause mortality, MI, stroke and T2DM onset was identified for each patient. Separate patient cohorts were created for each outcome because complete patient data were not available. Except for the T2DM cohort, each of these cohorts were then subdivided into diabetic and non-diabetic cohorts (only including patients who were diabetic

or non-diabetic for the entire follow-up), resulting in seven cohorts for which TTE were estimated. Covariates included in the Ara et al.[58] model include baseline age; sex; use of aspirin, statins, blood-pressure lowering treatment; and smoking status. Diabetic cohorts also included a covariate dummy for insulin use. Only baseline BMI was used in TTE analysis. Weibull models were fitted to estimate TTE, with the Weibull scale parameter depending on each of the covariates, irrespective of statistical significance, and higher-order polynomial terms of BMI and age, based on significance at the 5% level. The Weibull shape parameter was only allowed to depend on a subset of prepared covariates, based on significance at the 5% level.

The company's model uses these TTE cohorts to inform the major clinical event estimates. However, general population data are used to inform all-cause mortality because only less than 10% of patients in the diabetic all-cause mortality cohort had follow-up beyond this point. Beyond 15 years, Weibull TTE estimates for MI, stroke and T2DM onset are applied over the company's model time horizon for obesity-related non-fatal events because it was not clear to the company what alternative assumptions were used in the models by Ara et al.[58] and data from the GPRD cohort were sparse.

To inform 2. the BMI trajectory of patients throughout their lives, Ara et al.[58] used multilevel modelling of the repeated measures of BMI in the GPRD cohort, with age as the timescale. Patients with BMI below 25kg/m^2 at any time were excluded from the analysis. Exploratory trajectory plots from random patients were used to inform model specification, before applying multilevel models. The model was adjusted for age and sex and the interaction between age and sex.

ERG comment: The ERG considered it appropriate to use risk equations for obesity-related events and the natural history model of BMI as reported in Ara et al.[58] The ERG is, however, concerned that the estimation of obesity-related events is based on a patient population that has a lower BMI (based on the Ara et al. BMI natural history model) than that of the population represented on the COR trial programme.

5.2.7 Adverse events

The company states that it was unable to make trial data comparisons between AE associated with NB32 and orlistat because details from clinical literature and regulatory documents on orlistat were insufficient. The company quotes the opinion of one clinical expert as *"NB32 patients have a HRQoL benefit over orlistat patients as a result of AE differences"*.[1] The company also refers to AEs in the lower digestive tract that can be particularly unpleasant for patients, referring to their own data on file. Lastly, the company claims that, *"while no NB32-related deaths were observed across the COR trial programme and the NB-CVOT and IGNITE studies, and orlistat mortality risk from increased liver reaction risk cannot be ruled out based on clinical study data"*.[1] As a result of quality of life implications of AE being poorly understood (especially in relation to whether the incidence of some AEs is treatment- or condition-related), the company only considers costs of AEs. The company adds that not accounting for HRQoL impact of AEs is conservative: 1) in comparison to orlistat, considering the relative AE profiles (for which no evidence was provided); and, 2) in comparison to standard management, for which the company claims that although the AE profile associated with NB32 together with standard management is less good than that of standard management alone, the treatment effectiveness HRQoL benefits outweigh any negative NB32-related AE effects on HRQoL. Furthermore, the company notes that in clinical trials and practice, treatment-related AEs are generally quickly resolved, with only short-term effects upon HRQoL.

The AE data used in the model are derived from the COR trial programme AE incidence data.

ERG comment: The ERG considers the company's claim that not accounting for HRQoL impact of AEs in the economic model is conservative as highly questionable. No systematic overview of evidence was provided that showed that the AE profile of orlistat was indeed worse than that of NB32. With regards to the comparative AE profile of NB32 vs. SM, it is clearly stated in the CS that the NB32 AE profile shows a higher incidence of AEs in the gastro-intestinal tract and nausea than that of SM.[1] The ERG does not consider the company's argument that these need not be reflected because treatment benefits outweigh any negative NB32-related AE effects on HRQoL a valid argument because the HRQoL effects of NB32 are captured in the model and AE effects are not. The ERG wishes to highlight that this omission leads to a downward bias in the ICER of NB32 compared with standard management. Upon request, the company provided a scenario analysis in their response to clarification question B13, in which *"pragmatic application of on-treatment disutilities has been provided",*[9] assuming all AEs to be associated with a utility decrement of 0.05 for the duration of one week. This analysis increased the company's base-case ICERs against orlistat and SM by £188 and £87 per QALY gained, respectively. The ERG was satisfied that the impact of AEs on HRQoL was likely to be small.

5.2.8 Health-related quality of life

The company uses EQ-5D data from the literature to inform HRQoL in the economic model. This was because in the COR trial programme, only disease-specific QoL data were collected in all but the COR-II study, in which the SF-36 questionnaire was also administered. In the COR trials, HRQoL was assessed using the IWQOL-Lite questionnaire, which assesses the impact of weight on quality of life in the five domains of physical function, self-esteem, sexual life, public distress, and work.[1] However, according to the company, the requirement for a generic measure of HRQoL, the frequency of completion and limited follow-up of the COR trials limited the usefulness of these data for the purposes of the economic model.

The company therefore performed a systematic search for HRQoL studies and, after screening and eligibility assessment, 49 publications were identified from which a total of 39 studies were included in the review. Some of these studies examined the relationship between BMI and EQ-5D utility; others the relationship between weight-related comorbidities and utility. However, the inability to explain the impact of both weight and weight-related comorbidities on utility, limited the usefulness of most included studies. The company therefore explored the use of utility estimates derived from historic HSE patient EQ-5D data from Ara et al.[58] but discarded this option due to inconsistencies in the report.

The company used the PHE weight management economic assessment tool, which was identified through the review. It uses results from regression analysis of individual-level EQ-5D data drawn from HSE from 2011 to 2013. The model includes explanatory variables for BMI, age, gender, and obesity-related conditions (stroke, MI, cancer and T2DM). Both, Tobit model estimates and Ordinary Least Squares regression model estimates are presented. In the company's base-case, the Tobit model utility estimates are used, the OLS estimates are explored in a company's scenario analysis. No justification was provided for the preference of the Tobit over the OLS model. The company presented the relationships between BMI and related disease in an overview presented here in Table 5.8.

As stated in Section 5.2.7, adverse events were not assigned any dis-utilities in the economic model.

ERG comment: The PHE utility regression model was found by searching the NICE and PHE websites. The ERG is concerned that the regression model that informs the utility estimates does not appear to be published in a peer-reviewed journal. As a consequence, given the limited amount of details, the validity of these regression models to estimate utility values can therefore not be assessed by the ERG. Furthermore, the ERG was concerned that the presentation of the regression model used

to estimate patients' utilities did not include checking of the face validity associated with the health state utilities. The company, in their response to clarification question B11[9] provided a summary of example patient utilities, shown in Table 5.9 and compared these with published general population utilities.[69] The company noted that the utility values predicted by the Tobit model for the healthy population resembled the ones from the general UK population and that the remainder of the predicted utilities lay below these (Table 5.9), demonstrating face validity. The ERG questioned the company's preference for the Tobit model but was satisfied by the company's response to clarification question B11.b that Tobit models are generally more appropriate for the modelling of utilities than OLS models, particularly because the alternative OLS models disregard the lower and upper bounds commonly used for the estimation of utilities. The impact of AEs was not incorporated in the model, see Section 5.2.8.

Table 5.8: Public Health England weight management economic assessment tool v2 HSE EQ-5D data analysis

Covariate	Coeff.	Variance-covariance matrix									
		BMI	BMI^2	BMI^3	Age	Female	Stroke	MI	Cancer	T2DM	Const.
Tobit Model Estimates[a]											
BMI	0.05911	0.00008									
BMI^2	-0.00175	0.00000	0.00000								
BMI^3	0.00001	0.00000	0.00000	0.00000							
Age	-0.00440	0.00000	0.00000	0.00000	0.00000						
Female	-0.04054	0.00000	0.00000	0.00000	0.00000	0.00002					
Stroke	-0.18280	0.00001	0.00000	0.00000	0.00000	0.00001	0.00059				
MI	-0.16122	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00048			
Cancer	-0.16403	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00003	-0.00003	0.00028		
T2DM	-0.11093	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00002	0.00001	0.00001	0.00012	
Constant	0.67263	-0.00084	0.00002	0.00000	0.00000	-0.00008	-0.00010	0.00006	0.00002	-0.00001	0.00940
Ordinary Least Squares Regression Estimates											
BMI	0.03293	0.00003									
BMI^2	-0.00094	0.00000	0.00000								
BMI^3	0.00001	0.00000	0.00000	0.00000							
Age	-0.00219	0.00000	0.00000	0.00000	0.00000						
Female	-0.02258	0.00000	0.00000	0.00000	0.00000	0.00001					
Stroke	-0.12652	0.00000	0.00000	0.00000	0.00000	0.00000	0.00044				
MI	-0.11931	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00035			
Cancer	-0.10944	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00002	-0.00001	0.00017		
T2DM	-0.07800	0.00000	0.00000	0.00000	0.00000	0.00000	-0.00001	0.00001	0.00000	0.00007	
Constant	0.65792	-0.00028	0.00001	0.00000	0.00000	-0.00002	-0.00004	0.00001	-0.00002	-0.00002	0.00311

BMI, body mass index; EQ-5D, EuroQol 5-dimensions; HSE, Health Survey for England; MI, myocardial infarction; OLS, ordinary least squares; T2DM, Type II diabetes mellitus.

Notes: [a], Censoring limits were -0.594 and 1; sigma 0.33898 (standard error 0.00365) (both to 5 decimal places).

Table 5.9: Summary of utilities estimated with the Tobit model compared with general population utility estimates

Patient characteristics	Male	Δ	Female	Δ
General population, 30 years	0.93		0.91	
Healthy, 30 years, BMI = 27	0.92	-0.01	0.90	-0.01
Diabetic, 30 years, BMI = 27	0.87	-0.06	0.85	-0.06
History of MI, 30 years, BMI = 27	0.85	-0.08	0.82	-0.09
History of stroke, 30 years, BMI = 27	0.83	-0.10	0.81	-0.10
History of MI and diabetic, 30 years, BMI = 27	0.78	-0.15	0.75	-0.16
History of stroke and diabetic, 30 years, BMI = 27	0.76	-0.17	0.73	-0.18
Healthy, 30 years, BMI = 35	0.88	-0.05	0.86	-0.05
Diabetic, 30 years, BMI = 35	0.83	-0.10	0.80	-0.11
History of MI, 30 years, BMI = 35	0.80	-0.13	0.77	-0.14
History of stroke, 30 years, BMI = 35	0.78	-0.15	0.75	-0.16
History of MI and diabetic, 30 years, BMI = 35	0.72	-0.21	0.69	-0.22
History of stroke and diabetic, 30 years, BMI = 35	0.70	-0.23	0.67	-0.24
General population, 50 years	0.88		0.86	
Healthy, 50 years, BMI = 27	0.88	0.00	0.86	0.00
Healthy, 50 years, BMI = 35	0.84	-0.04	0.82	-0.04
General population, 70 years	0.79		0.77	
Healthy, 70 years, BMI = 27	0.84	0.05	0.81	0.04
Diabetic, 70 years, BMI = 27	0.77	-0.02	0.74	-0.03
History of MI, 70 years, BMI = 27	0.73	-0.06	0.70	-0.07
History of stroke, 70 years, BMI = 27	0.72	-0.07	0.69	-0.08
History of MI and diabetic, 70 years, BMI = 27	0.65	-0.14	0.61	-0.16
History of stroke and diabetic, 70 years, BMI = 27	0.63	-0.16	0.59	-0.18
Healthy, 70 years, BMI = 35	0.79	0.00	0.76	-0.01
Diabetic, 70 years, BMI = 35	0.71	-0.08	0.68	-0.09
History of MI, 70 years, BMI = 35	0.67	-0.12	0.63	-0.14
History of stroke, 70 years, BMI = 35	0.65	-0.14	0.62	-0.15
History of MI and diabetic, 70 years, BMI = 35	0.57	-0.22	0.54	-0.23
History of stroke and diabetic, 70 years, BMI = 35	0.55	-0.24	0.52	-0.25

BMI, body mass index; MI, myocardial infarction.

5.2.9 Resources and costs

Costs in the model consisted of drug acquisition costs, non-drug costs related to standard management (applicable to all treatments considered), obesity-related comorbidity costs and adverse event costs.

Resource identification, measurement and valuation studies

In CS Appendix 17, the PRISMA flow diagram of studies identified for the cost and resource use review is presented. In total, 1,515 citations were identified through database searching, bibliographic searching and from conference proceedings. After screening and eligibility assessment, 22 publications were included in the review, which represented 20 unique studies (10 cost studies, four resource use studies, four resource use and cost studies and two cost effectiveness studies). A tabular summary of the characteristics of each included study is provided in CS Appendix 17.

The company stated that the level of reporting was generally poor across studies, to the extent that it was difficult to elicit useful resource use estimates for this analysis. A notable exception to this was the study by Ara et al.[58] Hence the company used this study to inform healthcare resource use assumptions.

Intervention and comparators drug acquisition costs

Table 5.10 summarises the drug acquisition costs for NB32 (8mg naltrexone/90mg bupropion) and orlistat. NB32 is associated with a four week titration period over which the dosage increases from one tablet daily (week 0), via two tablets daily (week 1) and three tablets daily (week 2) to four tablets per day (week 3 onwards).[40] The dosage for orlistat is three capsules daily (without titration period). As both NB32 and orlistat are oral medicines, it is anticipated that there are no costs associated with their administration.

There are no drug costs associated with standard management.

Table 5.10: Drug acquisition costs

Treatment	Pack size	Cost per pack	Cost per tablet	Source
NB32	112	£73.00	£0.65	List price submitted to the Department of Health
ORL	84	£18.44	£0.22	MIMS[70a]

Source: CS Table 62

Abbreviations: MIMS, Monthly Index of Medical Specialities; NB32, naltrexone 32mg plus bupropion; ORL, orlistat.

[a]The company argued that evidence shows that branded version of orlistat (Xenical) accounted for less than 1% of the total prescription items for orlistat. Hence, costs for Xenical are not included by the company.

Standard management costs

The non-drug resource use items comprising standard management in the model are GP visits, nurse visits and blood tests. The unit price of a GP and nurse visit were assumed to be £44.00 and £14.47 respectively (PSSRU (2015)[71]) while this was £3.01 for the costs of a blood test (NHS reference costs (2015) – Code DAPS05[72]). The resource use (i.e. expected frequencies) were estimated based on a combination of reporting in the COR studies,[28] the publication by Ara et al.[58] and UK clinical expert opinion. Moreover, the non-drug resource use and costs related to standard management were assumed different for the first 56 weeks and thereafter (see Table 5.11).

The company stated that, according to clinical opinion (JW),[64] non-drug resource use received alongside NB32 in the COR-I and COR-II clinical trials is a good reflection of the average diet and exercise regimens prescribed for obese and overweight patients in the UK. Therefore, the company included five GP visits during the first year consistently with these trials (though the timing of the 12/16 weeks GP visit for the response assessment differed between treatments). An additional GP visit was added for the reassessment (of the 5% weight loss) at 56 weeks for NB32 and 52 weeks for orlistat and standard management. Moreover, in line with the study by Ara et al.,[58] the company assumed monthly visits to

a healthcare professional for weight management (i.e. a GP or nurse visit at least every four weeks) and blood tests at baseline and three months for patients on active treatments (i.e. NB32 or orlistat). In addition, the company assumed, based on clinical opinion (JW), that all weight management patients would have an annual blood test to monitor blood glucose levels (either at week 52 or 56).

From week 60 onwards it is assumed that patients would continue to have nurse visits every four weeks.

Based on Table 5.11, it can be calculated that the costs of standard management, during the first 56 weeks, are £403.22 for standard management adjunct to NB32 and orlistat, while this is £397.21 for standard management alone. The company stated that this is in line with clinical opinion, as patients receiving standard management alone would incur approximately the same non-drug resource use costs as patients receiving NB32 or orlistat alongside standard management (excluding additional blood tests for patients receiving adjunctive therapy). After the first year, the costs of standard management are £14.47 (one nurse visit) every four weeks, independent of the treatment.

Table 5.11: Non-drug resource use related to standard management

Time (weeks)	NB32			ORL			SM		
	GP[a]	Nurse[a]	Blood[a]	GP[a]	Nurse[a]	Blood[a]	GP[a]	Nurse[a]	Blood[a]
0	1	0	1	1	0	1	1	0	0
4	0	1	0	0	1	0	0	1	0
8	0	1	0	0	1	0	0	1	0
12	0	1	0	1	0	1	1	0	0
16	1	0	1	0	1	0	0	1	0
20	0	1	0	0	1	0	0	1	0
24	1	0	0	1	0	0	1	0	0
28	0	1	0	0	1	0	0	1	0
32	0	1	0	0	1	0	0	1	0
36	1	0	0	1	0	0	1	0	0
40	0	1	0	0	1	0	0	1	0
44	0	1	0	0	1	0	0	1	0
48	1	0	0	1	0	0	1	0	0
52	0	1	0	1	0	1	1	0	1
56	1	0	1	0	1	0	0	1	0
60+[b]	0	1	0	0	1	0	0	1	0

Source: CS Table 66

Abbreviations: GP, general practitioner; NB, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

[a]The costs of a GP visit (11.7 minutes) were £44.00 (PSSRU (2015)[71]), the costs of a nurse visit (15.5 minutes) were £14.47 (PSSRU (2015)[71]) and the costs of a blood test were £3.01 (NHS reference costs (2015) – Code DAPS05[72])

[b]These frequencies apply from Week 60 every 4 weeks while patients are still receiving treatment.

Obesity-related comorbidity costs

Costs associated with obesity-related comorbidities were retrieved from the literature (mainly from the literature review performed by Ara et al.[58] inflated from 2009 levels to 2015 levels[71]) and adapted

following UK clinical expert consultation (Table 5.12). Based on clinical opinion (JW),[64] it was assumed that the NHS costs associated with MI, stroke and T2DM are additive.

Table 5.12: Obesity-related comorbidity costs

Category	Cost	Source	Description (of primary) source
MI (Year 1)	£4,210.75	Literature review by Ara et al.[58a]	Economic evaluation of early high-dose lipid lowering therapy to avoid cardiac events, which used bottom-up costing methods and considered hospitalisation, procedural, medical resource use and drug costs.[73]
MI (Year 1+)	£345.91		
Fatal MI[b]	£1,390.80		HTA report evaluating the cost effectiveness of glycoprotein IIb/IIIa antagonists in non-ST elevation acute coronary syndrome.[74]
Stroke (Year 1)	£9,482.78		UK cost-of-illness model that aimed to estimate stroke-related costs over a 5-year period based on data from a randomised study comparing alternative strategies of stroke care.[75]
Stroke (Year 1+)	£2,664.16		
Fatal stroke[b]	£8,671.94		
T2DM (Year 1)	£347.57	Diabetes UK[76a]	It is not clear whether Ara et al.[58] incorporated T2DM costs after the first year of onset (the company interpreted the costs from Ara et al.[58] as the costs for the first year). Hence, it seemed inappropriate to the company to use the costs from Ara et al.[58] each year. Therefore, a more recent report summarised by Diabetes UK was used. This report estimated monitoring and medication costs to be between £300 and £370 per patient per annum. These costs based on a mix of Type 1 and Type 2 diabetic patients, were used in the absence of specific Type 2 data. However, it should be noted that Type 1 diabetics make up a small minority of cases.[77] An average of these two estimates (£335; at 2012 price level) was used in the model.
T2DM (Year 1+)	£347.57		

Source: CS Table 67 and CS section 5.5.3

Abbreviations: T2DM, Type 2 diabetes mellitus; MI, myocardial infarction;

[a]Costs inflated to 2015 levels using Personal Social Services Research Unit Hospital and Community Health Services inflation indices[71]

[b]Ara et al.[58] included a cost upon death, if the death was caused by MI or stroke. The figure for cardiovascular disease mortality as a proportion of overall mortality (31%) was taken from WHO 2016 data.[78] Of the deaths attributable to cardiovascular disease, the proportions of deaths caused by MI (43.1%), stroke (32.9%) or other causes (24.0%) were taken from WHO 2004 data.[79] From this information, mortality related to MI and stroke, as a proportion of overall mortality, was calculated as 13.4% and 10.2%, respectively.

Adverse event unit costs and resource use

AE rates for patients on NB32 and standard management were calculated based on the COR-I trial (Table 5.13, as the largest trial in the COR trial programme. Costs were considered for AEs that occurred in at least 5% of patients (either treatment arm). This threshold was selected to reflect the British National Formulary criteria of all very common (> 1 in 10) and the majority of common (1 in 100 to 1 in 10) AEs.[80] The company assumed that AEs are treated solely within primary care and costed £44.00, representing a single GP visit. This resulted in weekly AE costs, during treatment of £1.69 and £0.81 for NB32 and standard management, respectively.

According to the company, the level of reporting of AE data for orlistat (in the studies identified in Section 4.10 as well as EMA regulatory documents) was not sufficient to compare AE incidence in orlistat patients accurately to that in NB32 patients. Therefore, the company assumes the same weekly AE costs for patients treated with orlistat as calculated for NB32. The company indicated that it expected the safety profile of NB32 to be non-inferior versus orlistat and hence that this assumption is likely to be conservative.

Table 5.13: COR-I trial adverse event occurrences and rates

Adverse event	NB32 (total N=573)			Standard management (total N=569)		
	N	Probability (within study)	Instantaneous rate	N	Probability (within study)	Instantaneous rate
Anxiety	9	0.0157	0.00045	12	0.0211	0.00059
Constipation	90	0.157	0.00481	32	0.0562	0.00161
Depression	3	0.00524	0.00015	6	0.0105	0.00029
Diarrhoea	26	0.0454	0.00131	28	0.0492	0.0014
Dizziness	54	0.0942	0.00279	15	0.0264	0.00074
Dry mouth	43	0.075	0.0022	11	0.0193	0.00054
Headache	79	0.138	0.00418	53	0.0931	0.00271
Hot flush	30	0.0524	0.00151	7	0.0123	0.00034
Insomnia	43	0.075	0.0022	29	0.051	0.00145
Nasopharyngitis	29	0.0506	0.00146	31	0.0545	0.00155
Nausea	171	0.298	0.00998	30	0.0527	0.0015
Sinusitis	30	0.0524	0.00151	34	0.0598	0.00171
Upper respiratory tract infection	57	0.0995	0.00295	64	0.113	0.00331
Vomiting	56	0.0977	0.0029	14	0.0246	0.00069
Source: CS Tables 69 and 70						

ERG comment: The ERG considered it plausible to use Ara et al.[58] (identified in the review) to inform healthcare resource use assumptions. Regarding the costs of standard management, it is unclear to the ERG why the company added a GP visit for the 52 week assessment for patients receiving standard management only (in addition to the five GP visits during the first year which was considered to be reflective of UK clinical practice). Therefore, the ERG removed this GP visit for patients receiving standard management only.

Drug wastage associated with NB32 was not considered in the base-case model. However, when considering this in a scenario analysis (response to clarification question B15), it is illustrated that not

considering drug wastage is not conservative (ICER compared with orlistat increased by £3,426). Given the unavailability of data, the ERG was not able to consider drug wastage in the ERG base-case model.

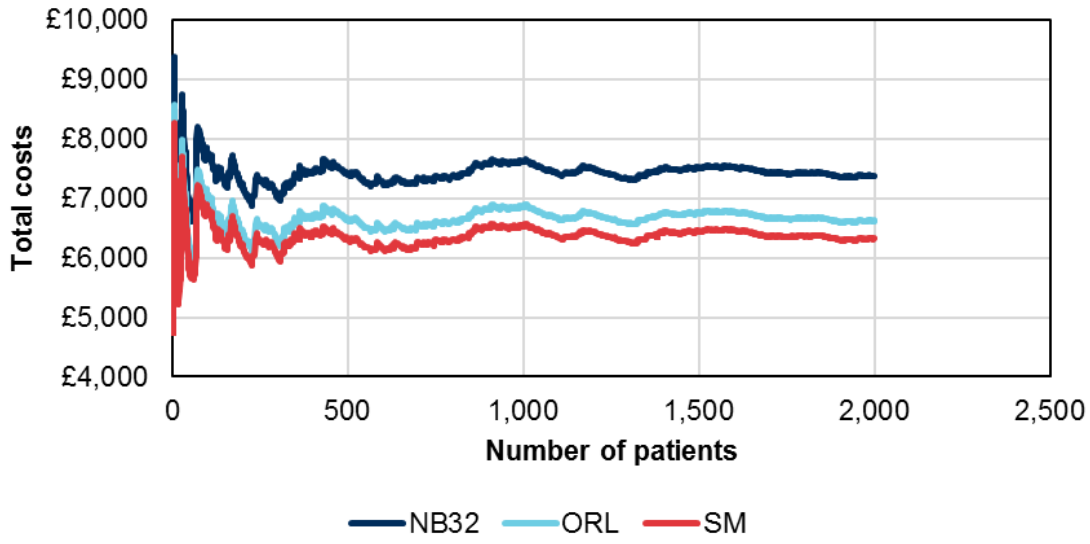
It was unclear to the ERG why the company considered it plausible to assume only a single GP visit for each adverse event. Assuming outpatient costs would increase the ICER of NB32 versus orlistat with £4,408 (CS Table 79). Moreover, it is unclear why the company expected the safety profile of NB32 to be non-inferior compared to the safety profile of orlistat. The company provided no systematic overview of evidence that showed that the AE profile of orlistat was indeed worse than that of NB32. There is no direct evidence comparing the two drugs and indirect treatment comparisons between the drugs focused on efficacy but not on safety outcomes. Therefore the company's assertion of the likely superiority of NB32 in relation to orlistat in terms of AE remains speculative. Therefore, it is questionable whether assuming the same AE costs for orlistat as calculated for NB32 is appropriate. Finally, it is unclear to the ERG why the company used the COR-I trial only to inform the rate of AE (e.g. why the COR-DM trial was not considered for T2DM specific AE rates).

5.2.10 Cost effectiveness results

Methodology for model analyses

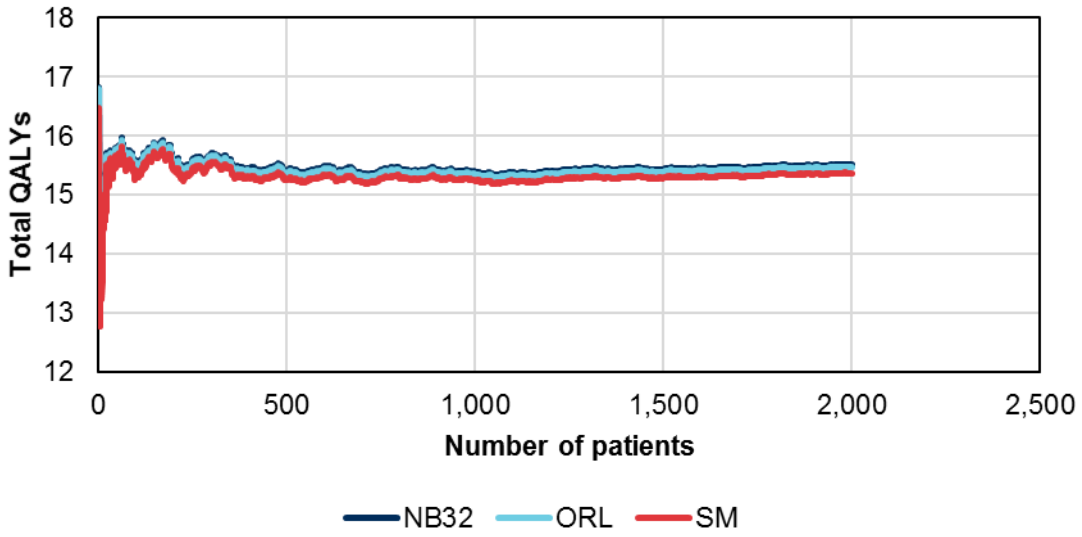
In order to obtain reliable results a sufficient number of individual randomly sampled random patient profiles need to be run such that the model results converge to a consistent value. In order to establish this number the company ran the model with 2,000 individual randomly sampled random patients and recorded total costs and total QALYs (Figures 5.10 and 5.11). Based on this exercise, the company decided that sampling 1,000 patients would a sufficient number to obtain a deterministic model result.

Figure 5.10: Diagnostic exercise – total costs



Source: CS, Figure 34
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

Figure 5.11: Diagnostic exercise – total QALYs



Source: CS, Figure 35
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

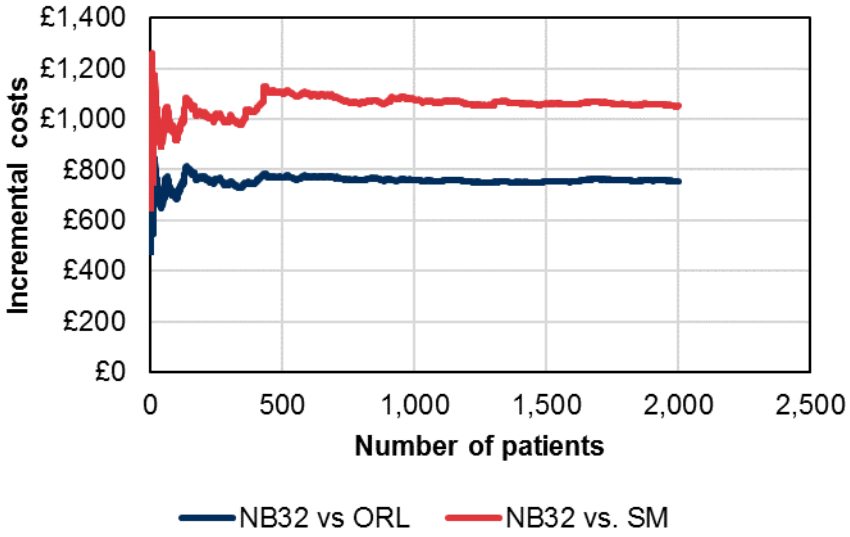
The probabilistic sensitivity analysis (PSA) took into account uncertainty surrounding cost estimates (except for drug cost and administration costs), utility estimates, BMI change in time, change in weight loss, and proportion of patients with response. Stochastic parameters not included in the PSA are TTD, administration costs (fixed zero), time until weight regain (fixed three years), and the probability of obesity related events (all cause mortality, non-fatal MI, non-fatal stroke, and onset of T2DM). Omitting to take into account uncertainty in the probabilities of all cause mortality and events was due to a lack of detail in the source.[58] An overview of the model inputs can be found in Appendix 18 of the CS.[1] For the PSA the company used 500 individual randomly sampled patients; the "*smallest number of patient profiles required after which model results appear to stabilise*". The number of PSA runs was chosen at 100, based on the run time required.

ERG comment: Ara et al.[58] used a cohort of 1,000,000 patients in their patient-level simulation and stated that, with a cohort size of 200,000 patients, there was still a small amount of variation in results, which stabilised after simulation of 400,000 patients. In contrast, a cohort of only 1,000 patients was used in the CS. The company provided two arguments to justify the lower number of sampled patients in comparison to Ara et al. First, the use of Excel instead of Simul8, which Ara et al. used, limited the number of sampled patients with regard to run-time. Second, the company argued that they "*were able to avoid the need to produce model results for a very large cohort (such as 1,000,000 patients) by controlling baseline characteristics for each model run. By controlling these characteristics, the only difference across patients was the treatment received.*"[9] The ERG finds this statement puzzling, as this is standard practice. In a patient-level simulation in each model run an identical individual randomly sampled patient should be evaluated for each of the comparators in the assessment.

The ERG asked the company to provide additional results from the diagnostic exercise to examine the minimum number of individual randomly sampled patients (incremental costs, QALYs and the incremental cost-effectiveness ratio (ICER)). The results indicate that for total and incremental QALYs 1,000 individual randomly sampled patients seems a sufficient number to obtain a reliable model result. For total and incremental costs and the ICER, the diagnostic exercise still shows fluctuations in results around 1,000 sampled patients (Figures 5.12, 5.13 and 5.14). The ERG ran the base case model with

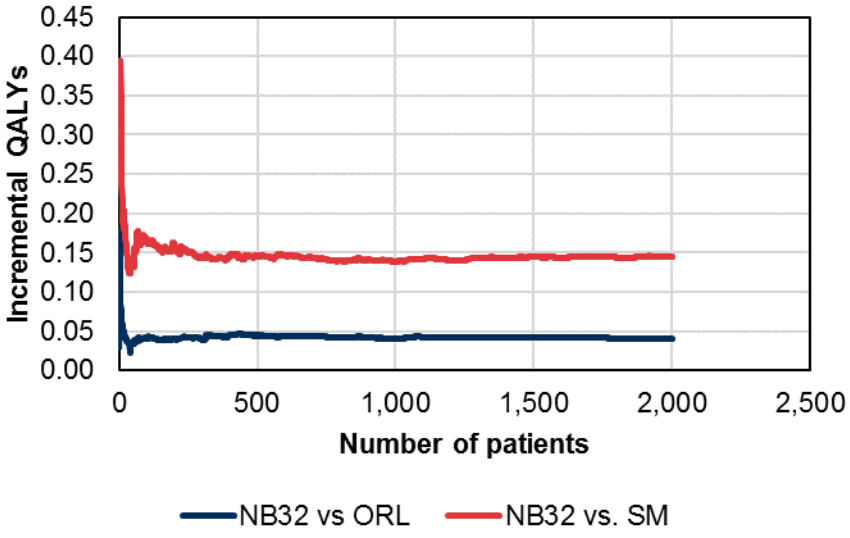
1,000 patients, which resulted in an ICER of NB32 versus orlistat ~£3,000 higher than the company's base case result. According to the ERG, the model should ideally be evaluated using at least 1,500 sampled patients.

Figure 5.12: Diagnostic exercise – incremental costs



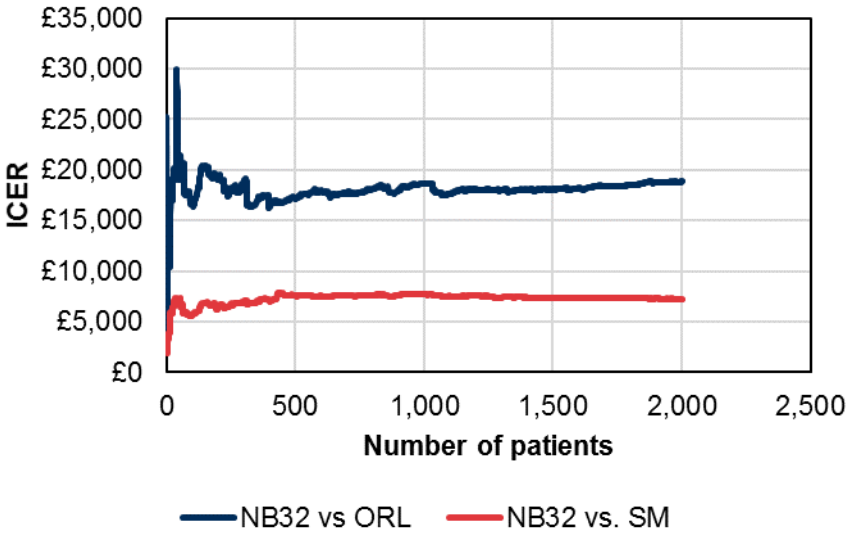
Source: Company response on clarification questions Figure 4
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.

Figure 5.13: Diagnostic exercise – incremental QALYs



Source: Company response on clarification questions Figure 5
Key: NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

Figure 5.14: Diagnostic exercise – ICER



Source: Company response on clarification questions Figure 6
Key: ICER, incremental cost-effectiveness ratio; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.

For the following stochastic parameters in the model uncertainty is not accounted for in the PSA: TTD, risk of obesity related events and the natural history of BMI model. The ERG asked the company to provide further clarification. The company stated that uncertainty of risk of obesity related events and natural history of BMI was not included in the PSA because Ara et al. 2012,[58] the source for key time to event equations, did not report variance-covariance matrices, and did not respond to email requests for these in time for the submission. According to the ERG, the company could have used standard errors, and/or have made assumptions to account for the uncertainty in these estimates in the PSA. The company also argued that "much of the key uncertainty regarding model results is structural and methodological, and based on the key conservative assumptions underpinning the analysis" and hence "the choice of the number of PSA iterations (be that 100, 100,000 or 100,000,000) does not demonstrate how conservative the structural model assumptions are.".[9] The ERG disagrees on that key methodological and structural assumptions are conservative (see Section 5.2.2). Moreover, the PSA is not only a method to show uncertainty around mean outcomes, but also the preferred method to obtain the mean outcomes.[9] Hence, if the PSA is flawed, so is the estimation of the mean outcomes of the model.

The PSA is performed with a smaller (500) number of individual randomly sampled patients. The ERG disagrees that this at this number of patients results appear to stabilise. As the figures of the diagnostic exercise show, at 500 patients the results for both QALYs and costs do not converge yet. As a result, the deterministic result on which the PSA runs are applied is unreliable. Another weakness of the PSA methodology the company used is the small number (100) of PSA runs. It is very unlikely this will result in a reliable probabilistic model estimate for an individual patient profile. Usually at least 1,000, and often much higher numbers of 5,000 to 10,000 PSA runs are required to obtain a reliable result. The ERG asked the company to provide a justification as to why the company believes that probabilistic sensitivity analysis using 100 simulations results in stable/plausible results. In their response, the company reiterated the choices made for an individual-level model *("better suited than a cohort-level approach to capture the chronic implications of both weight and weight-related health events in a heterogeneous group of overweight and obese patients"*), programmed in MS Excel *("applicable across various health technology assessment agencies internationally – some of which only consider*

models constructed within Microsoft Excel®. Microsoft Excel is also the preferred software package of NICE, and is typically considered more transparent than simulation models constructed in other software packages.").[9] The company further stated "*Regarding the number of PSA simulations, a trade-off between the number of patient profiles and the number of probabilistic draws was made. Given that within the PSA, 500 patient profiles are used for each PSA run, the number of PSA runs was chosen at 100. This number of PSA iterations was associated with a long run-time due to the limitations in processing power associated with Microsoft Excel*.".[9] Indeed, run-time of the model is relatively long. The ERG recorded run times between 450 and 600 hours of a PSA with 100 individual randomly sampled patients and 1,000 PSA runs. These numbers of sampled patients and PSA runs are still too low to obtain a reliable result. The ERG acknowledges that in any model study trade-offs are made between validity and reliability of the result and practical considerations. However, companies should provide a submission that is compliant to the NICE decision making process in which probabilistic model results are preferred, and the model is assessed by the ERG in a period of eight weeks. For this model, it was unfeasible for the ERG to perform an adequate assessment of the model's probabilistic results within the time frame of a NICE submission.

The ERG believes that the reliability of the probabilistic model results is severely compromised as a result of not accounting for uncertainty in some stochastic parameters, and instability due to too low a number of individual randomly sampled patients and too low a number of PSA runs.

Base-case model results

In the base-case deterministic analysis NB32 gains 0.0765 QALY versus standard management, and 0.0192 QALY versus orlistat. The incremental costs of NB32 are £1,044 versus standard management and £750 versus orlistat. The incremental cost-effectiveness ratio (ICER) of NB32 versus standard management is £13,647 per QALY. The estimated ICER versus orlistat is £32,084 per QALY. The latter ICER is also the ICER in a full incremental analysis. The probabilistic analysis shows a similar ICER for NB32 versus standard management (£13,958) and a higher ICER for NB32 versus orlistat (£36,084). See Tables 5.14 and 5.15 below.

Table 5.14: Base case deterministic results

Treatment	Total			Incremental			ICER (QALYs)	
	Costs	LYs[a]	QALYs	Costs	LYs[a]	QALYs	Versus baseline (SM)	Incremental
SM	£6,519	33.4768	15.3616					
ORL	£6,814	33.5151	15.4148	£294	0.0383	0.0531	£5,538	£5,538
NB32[b]	£7,563	33.5343	15.4381	£750	0.0192	0.0234	£13,647	£32,084

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.
Source: Table 72 CS
Note: [a] LYs are undiscounted, costs and QALYs are discounted. [b] The ICER of NB32 versus SM amounts to £13,647

Table 5.15: Probabilistic base case model results

Treatment	Total			Incremental			ICER (QALYs)	
	Costs	LYs[a]	QALYs	Costs	LYs[a]	QALYs	Versus baseline (SM)	Incremental
SM	£6,411	33.5673	15.3664					
ORL	£6,667	33.6128	15.4176	£256	0.0455	0.0512	£4,993	£4,993
NB32[b]	£7,409	33.6242	15.4379	£742	0.0115	0.0204	£13,936	£36,405

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.
Source: Adapted from Table 78 CS
Note: [a] LYs are undiscounted, costs and QALYs are discounted. [b] The ICER of NB32 versus SM amounts to £13,936

The company presented disaggregated results for costs. This shows that the cost differences between the comparators is caused by the treatment acquisition costs (Table 5.16)

Table 5.16: Summary of discounted costs by cost category

Technologies	Costs				
	Treatment acquisition	SM and CM	AEs	Death	Total
SM	£0	£5,982	£171	£367	£6,519
ORL	£238	£5,993	£216	£366	£6,814
NB32	£995	£5,983	£220	£366	£7,563

Key: AE, adverse event; CM, condition management; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; SM, standard management.
Source: CS Table 76

ERG comment: The deterministic total cost result should be interpreted with caution due a possible to small number of sampled patients to obtain a stable result. According to the NICE DSU guidance,[81] decision making should be based on probabilistic model results. However, in this submission, the PSA results are flawed, for reasons explained in the previous section.

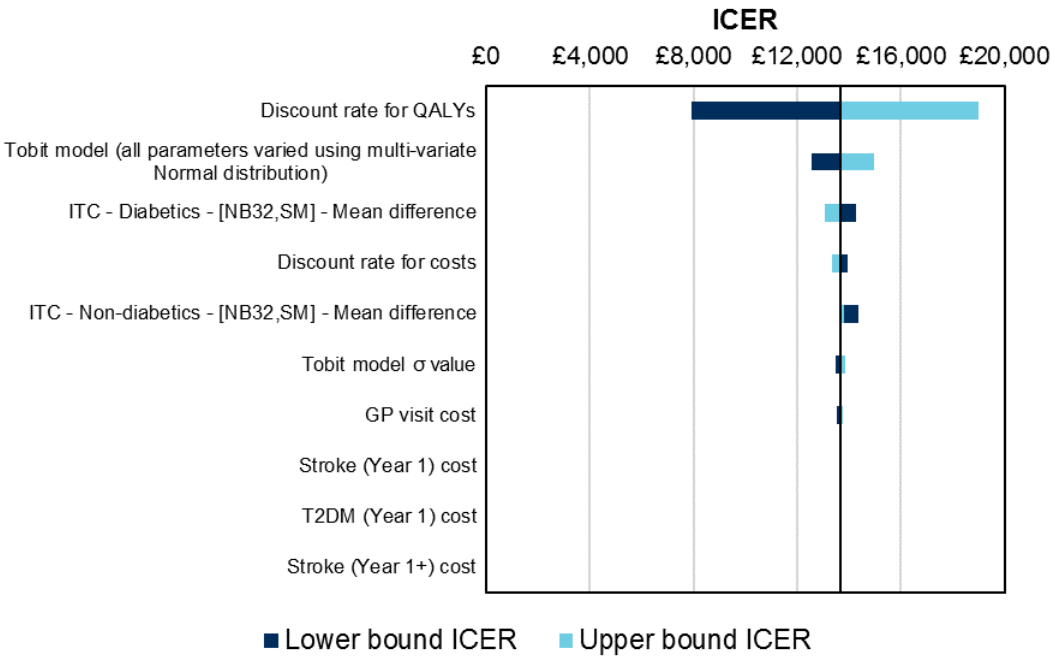
The ERG asked for more information on disaggregated outcomes of the model, such as costs associated with events, and time with events, but these were not provided by the company.

5.2.11 Sensitivity analyses

The company provided scatterplots and cost effectiveness acceptability curves based on the results of the PSA for NB32 versus standard management and NB32 versus orlistat separately. The cost effectiveness acceptability curves (CEAC) show that NB32 is associated with a 98% probability of being cost effective versus standard management and a 0% probability of being cost effective versus orlistat at a willingness to pay (WTP) threshold of £20,000 per QALY.

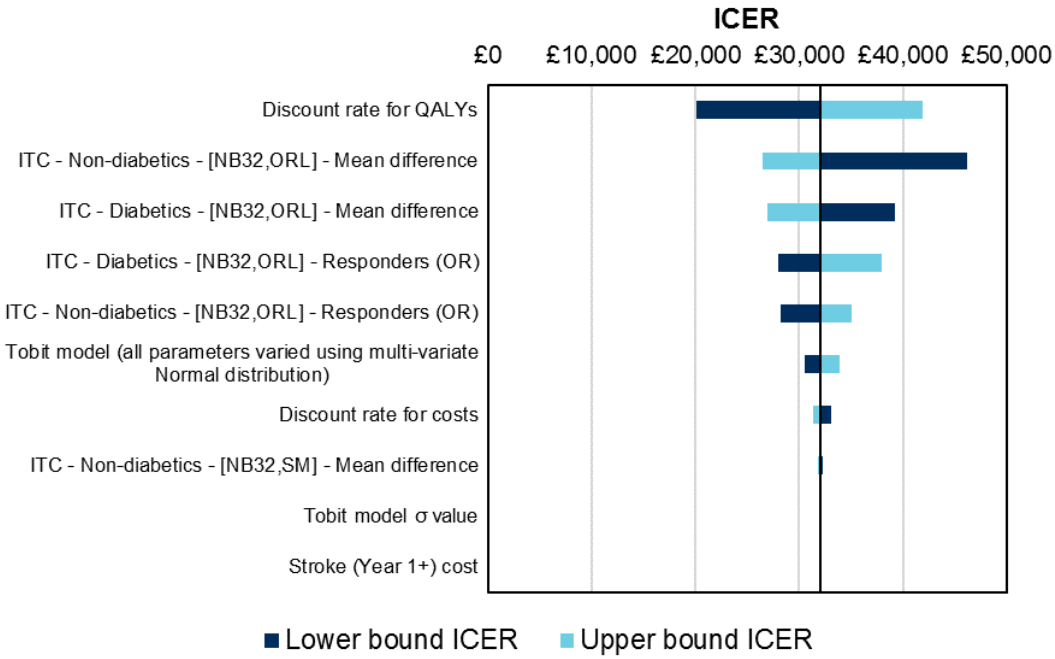
The deterministic sensitivity analyses performed by the company show that the most influential parameters are the parameters of the Tobit model for utilities and the discount rate for QALYs, as well as parameters related to the measures of relative efficacy from the ITC.

Figure 5.15: OWSA – NB32 versus SM



Key: GP, general practitioner; ICER, incremental cost-effectiveness ratio; ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; OLS, ordinary least squares; QALY, quality-adjusted life year; SM, standard management; T2DM, Type 2 diabetes mellitus. Source: CS Figure 42

Figure 5.16: OWSA – NB32 versus ORL



Key: GP, general practitioner; ICER, incremental cost-effectiveness ratio; ITC, indirect treatment comparison; NB32, naltrexone 32mg plus bupropion; OLS, ordinary least squares; OR, odds ratio; QALY, quality-adjusted life year; SM, standard management; T2DM, Type 2 diabetes mellitus. Source: CS Figure 43

Note: The eighth parameter (ITC – Non-diabetics – [NB32, SM] – Mean difference] is not an error. This parameter is featured within the outcome of the analysis as patients who discontinue treatment with orlistat may continue treatment with standard management alone.

The company performed scenario analyses on the following model aspects: the time period over which weight is regained, the cost of T2DM, the utility estimates, costs of AEs, discounting, and the time horizon. The most influential scenarios were shortening the time period for weight regain from three to two years (ICER £41,016), and shortening the time horizon form lifetime to 15 years (£53,514).

Table 5.17: Scenario analysis results

Scenario				ICERs	
				NB32 vs	
n	**Model setting**	**Base case**	**Scenario tested**	**ORL**	**SM**
0	Base case			£32,084	£13,647
1	Weight regain	3 years	2 years	£41,016	£14,113
2	Weight regain	3 years	5 years	£29,739	£11,880
3	Cost of T2DM	£347.57	£175.86 in Year 1 only	£36,096	£13,764
4	Utility model	Tobit	OLS	£36,771	£10,285
5	AE costs	All GP	All outpatient	£36,492	£15,130
6	Discounting	3.5% for costs & effects	1.5% for costs & effects	£28,323	£9,969
7	Time horizon	Lifetime	15 years	£53,514	£22,763

Key: AE, adverse event; GP, general practitioner; ICER, incremental cost-effectiveness ratio; LY, life year; NB32, naltrexone 32mg plus bupropion; OLS, ordinary least squares; QALY, quality-adjusted life year; SM, standard management; T2DM, Type 2 diabetes mellitus. Source: CS Table 70

The company performed subgroup analyses for patients with and without T2DM at baseline. The results showed that the ICER of NB32 versus orlistat it higher in patients who have T2DM at treatment initiation (£72,069), compared to patients who do not have T2DM at that moment (£28,298). The company warns that the results are uncertain because the data regarding comparisons of NB32 to orlistat in patients with T2DM are limited as shown, in Section 4.10 of the CS.[1]

Table 5.18: Base case results – T2DM patients at baseline only

T	Total			Incremental			ICER (QALYs)	
	Costs	**LYs[a]**	**QALYs**	**Costs**	**LYs[a]**	**QALYs**	**Versus baseline (SM)**	**Incremental**
SM	£10,199	32.7296	14.3707					
ORL	£10,496	32.7583	14.4295	£297	0.0287	0.0588	£5,059	£5,059
NB32	£11,216	32.7656	14.4395	£720	0.0073	0.0100	£14,797	£72,069

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; NB32, naltrexone 32mg plus bupropion; ORL, orlistat; QALY, quality-adjusted life year; SM, standard management.
Note: [a], LYs are undiscounted, costs and QALYs are discounted. Source CS Table 80

Table 5.19: Base case results – non-T2DM patients at baseline only

T	Total			Incremental			ICER (QALYs)	
	Costs	LYs[a]	QALYs	Costs	LYs[a]	QALYs	Versus baseline (SM)	Incremental
SM	£3,844	33.5497	15.7335					
ORL	£4,077	33.5854	15.7706	£233	0.0356	0.0371	£6,283	£6,283
NB32	£4,811	33.5944	15.7966	£734	0.0090	0.0259	£15,339	£28,291

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; 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. Source: CS Table 81

ERG comment: For reasons described in the previous paragraph, the ERG believes the PSA results are flawed. As a result the CEACs should be interpreted with extreme caution. The company did not perform deterministic sensitivity analyses on all parameters that are uncertain. Most notably, some parameters that were left out of the PSA were also not varied in deterministic sensitivity analyses, such as TTD, the probability of obesity related events, and the BMI natural history model. For instance the uncertainty around TTD, influencing both treatment effects and costs, is likely to significantly affect model results. The subgroup analyses with T2DM and non-T2DM patients should be interpreted with caution, because in these subgroup analyses the baseline characteristics (which impact obesity-related comorbidities and utility values) are independent on T2DM status. As stated in section 5.2.3 this leads to counter-intuitive patient profiles.

5.2.12 Model validation and face validity check

Face validity

The company attempted to achieve face validity by using advice from a clinical expert (JW). The company indicated that advice from this expert was used to inform and validate key clinical assumptions in the analysis.[64]

Internal validity

An economist not involved in model adaptation reviewed the model for coding errors, inconsistencies and the plausibility of inputs. This included examining known modelling errors, and questioning of the assumptions. In addition, in response to clarification question B19, the company stated that it used a checklist of basic validity checks (e.g. setting all costs to zero and ensuring the model outputs zero costs), sheet by sheet check of model logic (e.g. checking DICE equation logic), module by module check of VBA logic, validity assessment of outcomes (e.g. comparing available trial data with the outcomes of the model), and editorial checks (e.g. performing a spell check of model content).

External validity

The company compared the estimated LYs in the model (range: 33.48 to 33.53 for the three treatments) with UK life expectancy for the general population at the age of 47 years (range 34 to 37 for males and females respectively) and considered this to be a validation of the LYs estimated in the model.

Cross validity

The company considered that the total QALYs from their model are similar to those reported by Ara et al.,[58] for standard management (15.36 versus 15.13) and orlistat (15.41 versus 15.30).

ERG comment: The ERG considered the internal validity of the model (e.g. checking formula's in the DICE sheet, examining the implementation of TTD in the model, examining available intermediate outcomes). However, the ERG was unable to examine the internal validity of the model according to its usual standards. This was mainly a consequence of the time available to the ERG in relation to the time the model requires to run one single deterministic analysis, and the inability to examine intermediate outcomes. For instance, the nature of the model hampered the ERG's ability to do sensitivity analysis; extreme value analysis; trace analysis/analysis of intermediate outcomes which are recommended by the ISPOR taskforce on model transparency and validation.[82] Therefore, the ERG cannot guarantee that there are no modelling errors (in addition to the methodological flaws described below). In this light, the ERG considers it as troublesome that the company did not provide the results of the internal validation it performed (as requested in response to clarification question B19).[9]

The ERG wishes to highlight that it considers the model submitted by the company as unfit for purpose. The implementation in DICE resulted in extremely slow runtimes (6 hours on average, but with occasional model run times of 10 hours, depending on computer specifications) for the deterministic analysis only. It should also be noted that the model crashed on most of the computers that it was tried on.

One of the main validity issues or methodological flaws the ERG encountered was the lack of an updating event or integration of BMI over time. The average time between model entry and death was 33.5 years (median: 35.1 years; interquartile range 6.2 years - 56.7 years). The ERG calculated that on average patients in the model have seven events (excluding the start and end events), on average four of which occur during the first year (i.e. until the second assessment or if not applicable, due to treatment discontinuation during the first year, until date of treatment discontinuation). Hence after the first year, patients have on average only three events in 32.8 years, equalling to an average of one event per 10.6 years (median: 10.0 years; interquartile range 1.7 years - 23.9 years). This entails that BMI after the first year is only updated on average once every 10.6 years (implicitly assuming a stable BMI in the periods between events), while this should be updated at least annually to reflect the increasing BMI due to its correlation with age (as reflected in the natural history model predicting BMI over time). An alternative would be discrete integration of the BMI function. Apart from the continuous development of BMI not being reflected (and the impact on associated risks and utility values), the lack of model updating also affects other assumptions in the model. For example, the assumption regarding weight regain after treatment discontinuation for NB32 and orlistat was intended to reflect linear weight regain for a period of three years after which the BMI is obtained (predicted by the natural history model). However, if there is no event in this three year weight regain period, which is more likely than not (based on the average of one event per 10.6 years), the BMI estimated at the time of treatment discontinuation is maintained for this weight regain period of three years after which the weight is regained instantly. It should be noted that if the death event would be excluded from this calculation (also excluding 0.7% of the patients with death as the only event), the average time until one event increases to 17.2 years (median: 14.1 years; interquartile range 2.9 years - 47.9 years). The death event, logically, does not provide an intermediate update of BMI. Additionally, given that BMI is not updated at the stop adjunct, stop treatment and death events, the average period without BMI update presented above is an underestimation of the actual period without BMI update in the model. The ISPOR taskforce on DES[83] states that in case *"the likelihood of discrete events is a function of the value of a continuous measure (e.g., diabetic complications are a function of Hb A1c, or clinical presentation is a function of tumor size), as described in the model structure and design section"*, that *"time checks can be used to sample the likelihood of discrete events, conditional on the status of the continuous measure of disease progression (e.g., monthly time checks to update Hb A1c levels and define related probabilities of*

complications).” In the present model, the likelihood of events, as well as utility, is a function of BMI, which is predicted (by the natural history model) to change annually with increasing age. Hence, given the average time to event, it would have been recommended to incorporate ‘time checks’ (i.e. ‘update events’). Given that BMI is underestimated as a consequence of this methodological flaw, the utility values and the time to the events in the model are overestimated, likely inducing bias in favour of NB32. Moreover, assuming stable BMI for long periods of time also limits the face validity of the model.

Considering face validity, the ERG wishes to highlight that owing to the technical implementation of the model, it was difficult to assess the face validity of all parts of the model in the given time-frame. For example, the ERG identified one potential issue with the proportion of responders at secondary assessment. For both NB32 and orlistat this is directly determined by the estimated weight loss. Responders to these treatments at the primary assessment are therefore on average set to be responders at the one year assessment, too. As the company pointed out in their response to clarification question B5, not all patients are responders at the one year assessment in the model, due to the weight loss distribution being sampled from for each patient.[9] The weight loss distribution was a normal distribution with a mean weight loss of 11.7% and SD of 7.2% for NB32 and similarly for orlistat and SM based on ITC results, which means that only a small proportion of NB32 patients continuing on treatment after the primary assessment would be non-responders at one year (approximately 17%). When validating this assumption against the patient output, the ERG noted that proportions of non-responders at the one year assessment indicated a similar proportion of responders compared to the number of patients at baseline for patients treated with NB32, orlistat and only receiving standard management (57.3%, 55.6%, 56.6% of all patients at baseline, respectively). It is unclear how such similar proportions were obtained. These could be a result of the different events (treatment discontinuation, death, weight loss) but it was not possible to check whether these estimates truly exhibited face validity. It should therefore be considered whether simpler approaches (e.g. an individual-level state transition model) would have been more appropriate to reflect this decision problem, given the gain in transparency and given that it would have been possible to reflect the condition-specific events in such a model. A state-transition approach would potentially resolve most of the validity issues (e.g. the lacking updating event).

The ERG noted a small inconsistency in the implementation of the TTD estimation in the model. The TTD estimates for NB32 and orlistat were a result of 1. sampling from TTD KM estimates of patients up to the primary assessment (16 weeks for NB32 and 12 weeks for orlistat), 2. sampling from TTD KM estimates of patients up to the secondary assessment (56 weeks for NB32 and 52 weeks for orlistat), and 3. sampling from the TTD KM estimates for the remainder of time until there was no more available data from the NB-CVOT study (the maximum was a total of three years). The way that the sampling was done in each of these three steps was by using randomly generated numbers between 0 and 1, then finding the closest matching KM estimate (that is percentage of patients still on treatment) and then looking up the associated time to discontinuation. A vlookup function uses the random number and matches it to the largest value that it finds which is smaller than the random number, going through the lookup table from top to bottom (the table is sorted starting with the lowest percentage of patients still on treatment to the highest). The way the company sorted the values in lookup tables is different for NB32 and orlistat for the time period up to primary assessment – this results in very slightly smaller values of TTD for NB32 than for orlistat, but the impact of this inconsistency is expected to be minimal.

5.3 Exploratory and sensitivity analyses undertaken by the ERG

Table 5.20 summarises all main issues highlighted by the ERG in Section 5.2, indicates the expected direction of bias introduced by these issues and whether these are examined in any analyses/incorporated in the ERG base-case.

Table 5.20: Main ERG critique of company's submitted economic evaluation

Issue	Bias introduced[a]	ERG analyses (analysis number in section 5.3)	Addressed in analysis?
Model structure (section 5.2.2)			
• Weight regain assumptions deviated from those in Ara et al.[58] in that the company modelled weight regain towards the predicted BMI instead of the baseline BMI.	+	Base-case (7)	Response to clarification question B1; ICER (NB32 vs orlistat) increased by £1,536.
• Weight regain is not implemented linearly in the economic model.	+	Base-case (1)	
• 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.	-		
Population (section 5.2.3)			
• Baseline BMI is vastly underestimated in the economic model, hence overestimating utility and time to T2DM, cardiovascular events and death.	+	Base-case (4)	
• The proportions of current smokers, patients receiving anti-hypertensive medication and/or statins are most likely underestimated.	+	Base-case (5)	
• Counter-intuitive patient profiles are generated as correlations between patient characteristics are not incorporated.	+/-	Base-case (5)	
• 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.	+/-		
Interventions and comparators (section 5.2.4)			
• Behaviour interventions, bariatric surgery and re-treatment are not implemented.	+/-		
Treatment effectiveness and extrapolation (section 5.2.6)			
• The company used modified ITT analysis instead of ITT analysis for estimation of percentage of weight loss and response rates.	+	Base-case (2)	
• 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)	

Issue	Bias introduced[a]	ERG analyses (analysis number in section 5.3)	Addressed in analysis?
• An assumption is made that weight loss is equivalent for NB32 and orlistat at different times (16 weeks and 12 weeks, respectively).	+/-		
• 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)	
• 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)	
• 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.	+		
Adverse events (sections 5.2.7-5.2.9) • AE-related utility decrements were not included.	+		Response to clarification question B13; ICER (NB32 vs orlistat) increased by £188.
• 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.	+/-		
• Questionable whether the assumption of equivalent AE costs for NB32 and orlistat is conservative.	+/-		CS Table 79; using outpatient costs would increase the ICER of NB32 versus orlistat with £4,408
Health-related quality of life (section 5.2.8) • Use of PHE weight management economic assessment tool for derivation of utilities may not be appropriate.	+/-		
Resources and costs (section 5.2.9) • An unnecessary GP visit, related to response assessment, is incorporated for standard management.	+	Base-case (6)	

Issue	Bias introduced[a]	ERG analyses (analysis number in section 5.3)	Addressed in analysis?
• Assuming only a singly GP visit for each adverse event without plausible justification.	+		
• NB32 drug wastage was not considered in the model	+		Response to clarification question B15; ICER (NB32 vs orlistat) increased by £3,426.
Cost-effectiveness analyses (sections 5.2.10 and 5.2.11)			
• The number of simulated patients (1,000) is too low to provide stable results; the ICER varies substantially with each model run.	+/-		The ICER presented by the company was ~£3,000 lower than the one obtained by the ERG after re-running the deterministic analysis.
• The PSA does not appropriately reflect uncertainty surrounding the most important parameters (e.g. the uncertainty surrounding TTD, a key parameter in the model, was neglected).	+/-		Implementation of PSA violates best practices and hence should not be used for decision making. Moreover, the model run times are prohibitive to appropriately run PSA.
• The number of PSA simulations is restricted to 100, which is too low to appropriately reflect uncertainty	+/-		
Validation (section 5.2.12)			
• The lack of an updating event or discrete integration of the BMI function 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.	+		

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.

Based on all considerations from Section 5.2 (summarised in Table 5.20), the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016[84]):

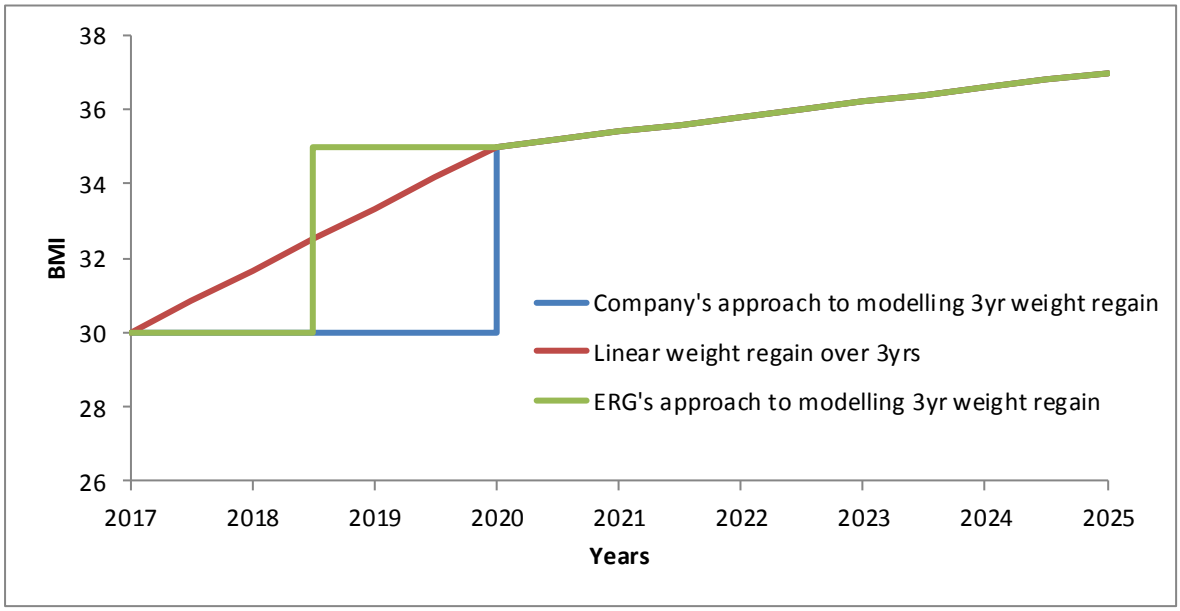
- Fixing errors (correcting the model were the company's submitted model was unequivocally wrong)
- Fixing violations (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (amending the model were the ERG considers that reasonable alternative assumptions are preferred)

Additionally, exploratory sensitivity analyses were performed by the ERG to examine the potential impact of alternative assumptions on the cost effectiveness estimates.

Fixing errors

1. Fixing errors consisted of using a weight regain period of 1.5 years after which weight is instantly regained, to reflect the three year linear weight regain assumption made by the company (see Figure 5.17 for an illustration of this with an example assuming treatment discontinuation at start of 2017). In the company base-case the weight is regained instantly after 3.0 years (instead of linearly over three years' time), see Sections 5.2.2 and 5.2.6 for more details. The ERG's approach first under-estimates BMI, then over-estimates it. In contrast to this, the company's approach under-estimates BMI for the whole duration of three years.

Figure 5.17: Illustrative example of ERG implementation of weight regain to reflect the three year linear weight regain assumption



Fixing violations

2. Using the ITT data instead of mITT data based on the COR-I and COR-DM trial only.
   The ERG considered the usage of ITT data from the COR-I and COR-DM trials as most appropriate (see Section 5.2.6 for more details).

3. Using a relative risk instead of mean difference to extrapolate the difference between treatments in change from baseline weight from the secondary to the primary assessment.

   The ERG calculated a relative risk for the difference between treatments in change from baseline weight based on the secondary assessment (which was based on the ITC). The relative risk from the secondary assessment (instead of using the mean difference as done in the CS) was applied to the change from baseline weight at the primary assessment (instead of using the mean difference as done in the CS), see Section 5.2.6 for more details.

4. The natural history model to predict BMI is calibrated to reflect the baseline BMI distribution as observed in the COR trial programme.

   The patient characteristics in the COR trial programme were considered a fair reflection of the typical patient group that would receive NB32 in UK NHS clinical practice. To maintain consistency between effectiveness estimates and the population in which these were derived, the ERG preferred to reflect the baseline BMI distribution as observed in the COR-I (for non T2DM patients) and COR-DM (for T2DM patients) trials in the economic model (see Section 5.2.3 for more details). The calibration was performed using a minimisation of sum of squared error terms that was operationalized using Solver in Excel in two steps, and separately for T2DM patients and non-T2DM patients:

   a. Calibrate the constants of the natural history model to predict BMI (calibrated to reflect a mean BMI of 36 kg/m2, as observed in the COR trials).

   b. Calibrate the variance of the constants, to calibrate the distribution over the BMI groups (calibrated based on proportions in BMI categories, using the sum of squared differences compared with the COR-I/COR-DM trials).

5. Adjust the baseline age (dependent on T2DM status), proportions of females (dependent on T2DM status), proportion of smokers, proportion receiving statins (dependent on T2DM status), proportion receiving anti-hypertensive medication (dependent on T2DM status) and proportion receiving aspirin.

   The ERG preferred to use baseline characteristics from the COR trial programme and stratified for T2DM status, if applicable (see Section 5.2.3 for more details).

6. Removal of GP visit for standard management.

   The GP visit for the 52 week assessment for patients receiving standard management only (in addition to the five GP visits during the first year, which was considered to be reflective of UK clinical practice) was removed (see Section 5.2.9 for more details).

Matters of judgment

7. Weight regain towards baseline BMI was assumed.

   The ERG noted that the company deviated from the assumption made by Ara et al.,[58] that patients would have regained weight to obtain the baseline BMI in three years and assumed instead that patients would have regained weight to obtain the predicted BMI in three years. The company did not provide justification for why this deviation was 'logical' and more plausible. To be consistent with Ara et al.,[58] the ERG preferred to assume weight regain to the baseline BMI in its base-case (see Section 5.2.2 for more details).

8. Remove linear scaling assumption for TTD of orlistat.

   The ERG believes that the linear scaling of TTD estimates for NB32 to obtain orlistat TTD may result in underestimating TTD for orlistat (see Section 5.2.6 for more details).

5.3.1 Deterministic ERG base-case

Given the flaws highlighted for the PSA, the ERG was restricted to doing a deterministic analysis using 1,000 patient profiles (as the maximum number of patient profiles was restricted to 1,000) to obtain the ERG base-case incorporating all abovementioned adjustments (see Table 5.21 for the BMI distribution sampled in the ERG base-case). The ERG did calculate the ERG base-case two times, each time based on different random numbers and a different set of sampled patients. The ERG base-case ICERs (deterministic) of NB32 compared with standard management and orlistat ranged between £9,813-£10,510 and £38,871-£45,694 per QALY gained respectively (see Table 6.1).

Table 5.21: BMI distribution in ERG base-case

	T2DM			No-T2DM		
	CS	**ERG**	**COR-DM**	**CS**	**ERG**	**COR-I**
BMI<30kg/m2	9%	0%	6%	8%	0%	2%
BMI ≥30 and ≤35 kg/m2	75%	35%	31%	75%	35%	38%
BMI ≥35 and <40 kg/m2	16%	62%	35%	18%	64%	37%
BMI ≥ 40 kg/m2	0%	2%	28%	0%	1%	23%

5.3.2 Additional exploratory analyses performed based on the ERG base-case

Additional sensitivity analyses were performed to examine the potential impact of the following alternative assumptions on the cost effectiveness estimates:

1. Using an instantaneous weight regain at the point of three years
2. Using a lower proportion (15%) of T2DM patients

The exploratory analyses indicated that using an instantaneous weight regain at the point of three years and a lower proportion (15%) of T2DM patients decreased the ICERs (Table 6.1).

5.3.3 Subgroup analyses performed based on the ERG base-case

Subgroup analyses were performed for patients with and without T2DM based on the ERG base-case. For patients with T2DM, NB32 was dominated by orlistat while the ICER versus standard management was £10,535 per QALY gained. In the subgroup without T2DM, NB32 compared with standard management and orlistat resulted in ICERs of £9,594 and £25,744 per QALY gained respectively (see Table 6.1).

5.4 Conclusions of the cost effectiveness section

The majority of the cost effectiveness searches in the CS were well documented and easily reproducible, and were carried out in line with the NICE guide to the methods of technology appraisal.

Reviewing the overall evidence, the ERG confirmed that there was no existing cost effectiveness model for NB32 for the current indication, and thus that development of a de novo model was necessary. The economic model described in the CS is considered by the ERG to meet the NICE reference case on most points. However, the analyses performed by the company were flawed (too low a number of sampled patient profiles, too low a number of PSA simulations and key parameters were not incorporated in the PSA) and deviated from the NICE reference case stating that probabilistic model

results are preferred. The company developed a de novo economic model using an individual-level approach, more specifically a discrete event simulation (DES). It was argued that an individual-level approach is better suited than a cohort-level approach to capture the chronic implications of both weight and weight-related health events in a heterogeneous group of overweight and obese patients. The DES model was implemented in Excel using the DICE principles and structure proposed by Caro.[63] In addition, the company used the economic evaluation by Ara et al.[58] (also an individual-level model) as a starting point, which is a Health Technology Appraisal (2012) comparing different pharmacological treatments for obesity. The model considered the following events:

- treatment discontinuation;
- development of T2DM;
- first cardiovascular event (either stroke or MI);
- second cardiovascular event (either stroke or MI) and;
- death.

The company base-case ICERs (deterministic) of NB32 compared with standard management and orlistat were £13,647 and £32,084 respectively. The deterministic sensitivity analyses performed by the company show that the most influential parameters are the parameters of the Tobit model for utilities and the discount rate for QALYs, as well as parameters related to the measures of relative efficacy from the ITC. These analyses, as well as the PSA performed by the company should be interpreted with extreme caution given the flaws highlighted above. Subgroup analyses performed by the company indicated that the ICERs (deterministic) of NB32 compared with standard management and orlistat were £14,797 and £72,069 per QALY gained respectively for T2DM patients and £15,339 and £28,291 per QALY gained respectively for non-T2DM patients.

The main issue with the company's model was its structure and its technical implementation which caused long run times (6 hours on average), and which caused the model to crash on multiple computers. This hampered the company's and the ERG's ability to perform an appropriate PSA and the ERG's ability to check the model's validity and perform further scenario analyses (other than those that were described in Section 5.3). It should be considered whether simpler approaches (e.g. an individual-level state transition model) would have been more appropriate to reflect this decision problem, given the gain in transparency and that it would have been possible to reflect the condition-specific events in such a model. An individual-level state transition approach would potentially resolve most of the validity issues (e.g. the fact that BMI was not accurately reflected at each time period).

Apart from that, numerous issues were identified by the ERG, the most important of which are summarised in Table 5.20. The ERG was able to adjust/correct some of these issues in its base-case. The ERG base-case ICERs (deterministic) of NB32 compared with standard management and orlistat ranged between £9,813-£10,510 and £38,871-£45,694 per QALY gained respectively. Subgroup analyses performed conditional on the ERG base-case, indicated that the ICERs (deterministic) of NB32 compared with standard management and orlistat were £10,535 per QALY gained and dominated respectively for T2DM patients and £9,594 and £25,744 per QALY gained respectively for non-T2DM patients. However, it should be noted that several issues remained unexplored (of which several were expected to be non-conservative, see Table 5.20) and thus the results should be interpreted in this context (i.e. with extreme caution). The interpretation and validity of the results are particularly hampered given that the company's model did underestimate TTD, did not incorporate behaviour modification interventions, bariatric surgery and re-treatment nor an updating event that was required to accurately reflect patients' expected quality of life and costs associated with resource use. As

discussed in Section 5.2.12, the fact that BMI development was not reflected in the model could significantly bias the results in favour of NB32.

The large variation around the ICERs when different random numbers and sampled patient profiles are used is of particular concern. In two different model runs of the ERG base-case, the ICER varied by as much as £7,000 per QALY gained. It is therefore the ERG's view that the company's model is of very limited value for the current decision problem and that results are to be interpreted with extreme caution.

In conclusion, given that the deterministic ERG base-case ICER of NB32 versus orlistat is estimated to range between £38,871 and £45,694 per QALY gained (based on different random numbers and different samples of patients), and the remaining issues/methodological flaws highlighted above, uncertainty around the cost effectiveness estimates of NB32 remains substantial.

6. IMPACT ON THE ICER OF ADDITIONAL CLINICAL AND ECONOMIC ANALYSES UNDERTAKEN BY THE ERG

In Section 5.3 the ERG base-case was presented, which was based on various changes compared to the company base-case. Table 6.1 shows the ERG replication of the company base-case, the ERG base-case, the exploratory analyses and subgroup analyses performed by the ERG (conditional on the ERG base-case). Appendix 1 contains technical details on the analyses performed by the ERG.

Table 6.1: ERG base-case, exploratory and subgroup analyses

	Technologies	Total costs	Total QALYs	NB32 Incremental costs	NB32 Incremental QALYs	NB32 ICER (£/QALY)
ERG base-case 1*	NB32	£7,017	15.21			
	Orlistat	£6,275	15.20	£742	0.02	£45,694
	SM	£5,964	15.11	£1,053	0.10	£10,510
ERG base-case 2*	NB32	£7,188	15.08			
	Orlistat	£6,455	15.06	£733	0.02	£38,871
	SM	£6,141	14.97	£1,047	0.11	£9,813
Company's base-case	NB32	£7,563	15.44			
	Orlistat	£6,814	15.41	£749	0.03	£32,084
	SM	£6,519	15.36	£1,044	0.08	£13,647
ERG replication of company's base-case	NB32	£6,948	15.36			
	Orlistat	£6,219	15.33	£729	0.02	£34,994
	SM	£5,974	15.29	£973	0.06	£15,568
Exploratory analyses						
1) Using instantaneous weight regain at 3 years	NB32	£7,048	15.19			
	Orlistat	£6,311	15.17	£737	0.02	£37,947
	SM	£6,007	15.09	£1,041	0.10	£10,021
2) Lower proportion (15%) of T2DM patients	NB32	£5,740	15.55			
	Orlistat	£4,992	15.53	£748	0.02	£28,687
	SM	£4,702	15.45	£1,038	0.10	£10,013
Subgroup analyses						
3) Subgroup non-T2DM patients	NB32	£4,603	15.77			
	Orlistat	£3,844	15.74	£759	0.03	£25,744
	SM	£3,565	15.66	£1,038	0.11	£9,594
4) Subgroup T2DM patients	NB32	£12,213	14.08			
	Orlistat	£11,527	14.09	£686	0.00	Dominated
	SM	£11,173	13.98	£1,040	0.10	£10,535

*These results are due to random variation between different model runs.

7. OVERALL CONCLUSIONS

7.1 Statement of principal findings

The four main trials comparing NB32 to placebo are of high quality. However there are a number of limitations when applying them to clinical practice. There are very little data on ethnic groups relevant to the UK (particularly people from Asia) within the NB32 trials, therefore it is not possible to make any firm conclusions for that group. There are very few overweight as opposed to obese participants in the trials. The majority of the participants in the NB32 trials are female. Trials do not measure weight loss beyond 56 weeks. The large dropout from the NB32 trials (up to 50%) is relevant to practice. The US setting may reflect a different patient profile and differing approaches to standard care than in a UK setting.

A comparison between NB32 (plus standard management) versus intensive behaviour modification is missing. Furthermore, comparisons between NB32 and orlistat are based on indirect comparisons only.

The company used modified ITT data from NB32 trials, but this is misleading. The mITT population in the NB32 trials is very different from mITT populations in the orlistat trials. In the NB32 trials, 21.9% of patients receiving NB32 were randomised but excluded from the analyses against 1.6% of patients receiving orlistat.

Comparison with orlistat may be biased in favour of NB32. NB32 trials were published in 2010 or later; most of the trials with orlistat were published before 2005, so caution should be exercised when making indirect comparisons; this is particularly true for conditions such as diabetes where background standard therapy (for glucose and lipids especially) may be very different now.

We have reproduced the company's indirect analyses comparing orlistat and NB32 using full ITT data from the NB32 trials and we have included a new analysis: an indirect comparison of NB32 plus intensive behaviour modification (COR-BMOD) versus orlistat plus intensive behaviour modification (XENDOS). The results show that the positive effects of NB32 when compared to orlistat have all disappeared. For the first outcome (≥5% reduction in weight at one year), there was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded using mITT data. In both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: OR = 1.09 (95% CrI: 0.87 to 1.36), ITT-BOCF: OR = 1.06 (95% CrI: 0.84 to 1.33). Moreover, although none of the differences are statistically significant, all results now favour orlistat.

For the second outcome (mean percentage weight change at one year), there was a statistically significant difference favouring NB32 over orlistat in the analyses where studies with T2DM patients were excluded using mITT data. In both ITT analyses there is no significant difference between NB32 and orlistat for studies with T2DM patients excluded (ITT-Imp: MD= -0.09 (95% CrI: -0.77 to 0.58), ITT-BOCF: MD = -0.54 (95% CrI: -1.21 to 0.12). Moreover, although most of the differences are not statistically significant, most results now favour orlistat.

The results of the indirect comparison of NB32 plus intensive behaviour modification versus orlistat plus intensive behaviour modification, using data from COR-BMOD versus XENDOS, show that both outcomes significantly favour orlistat over NB32 (≥5% reduction in weight at one year: OR 1.86 (95% CI: 1.30 to 2.66); mean percentage weight CFB at one year: MD -2.09 (95% CI: -3.53 to -0.65)).

Finally, we performed our preferred analyses, i.e. using full ITT data and no pooling of NB32 trials (using only COR-I ITT data for non-diabetics, instead of COR-I, COR-II and COR-BMOD combined).

The results for 'obese patients with T2DM' and 'intensive behaviour modification' are the same as before, but results for 'obese patients without T2DM' have changed considerably again, and are almost the same as in the company's original analyses. Both outcomes show no significant difference between NB32 and orlistat, but both favour NB32 in this subgroup.

The table below shows the main results for obese people with diabetes, obese people without diabetes and NB32 plus intensive behaviour modification versus orlistat plus intensive behaviour modification.

Table 7.1: Company results versus ERG results

Population		Company analyses (mITT data)*	Company analyses (ITT-BCFA data)**	ERG preferred analyses**
		Orlistat vs NB32	Orlistat vs NB32	Orlistat vs NB32
Obese people with T2DM				
≥5% reduction in weight at 1 year	OR	1.09 (0.63 to 1.88)	1.59 (0.89 to 2.79)	1.59 (0.89 to 2.79)
Mean % weight CFB at 1 year	MD	0.21 (-0.87 to 1.30)	-1.21 (-2.30 to -0.11)	-1.21 (-2.30 to -0.11)
Obese people without T2DM				
≥5% reduction in weight at 1 year	OR	0.77 (0.61 to 0.96)	1.06 (0.84 to 1.33)	0.61 (0.31 to 1.22)
Mean % weight CFB at 1 year	MD	1.13 (0.44 to 1.80)	-0.54 (-1.21 to 0.12)	1.11 (-0.39 to 2.63)
Intensive behaviour modification				
≥5% reduction in weight at 1 year	OR	1.22 (0.84 to 1.77)	1.86 (1.30 to 2.66)	1.86 (1.30 to 2.66)
Mean % weight CFB at 1 year	MD	-0.21 (-1.28 to 1.70)	-2.09 (-3.53to -0.65)	-2.09 (-3.53to -0.65)

Results are OR with 95% CI/CrI for ≥5% reduction in weight at 1 year and mean difference (MD) with 95% CI/CrI for mean % weight CFB at 1 year.
An OR less than one favours NB32 over orlistat and a CI including 1 is not significant. A MD of >0 favours NB32 over orlistat and indicates greater % weight reduction and a CI including 0 is not significant.
*) Bayesian NMA (OR, 95% CrI) using mITT data; **) Using the Bucher method for indirect comparisons and ITT-BCFA data.
FE = fixed effect; ITT-BCFA = all randomised patients with baseline-carried-forward analysis; MD = Mean Difference; mITT = modified intention-to-treat analysis; NB32 = naltrexone 32mg plus bupropion; OR = Odds Ratio; T2DM = Type 2 diabetes mellitus;

Which of the estimates of treatment effect is more applicable to clinical practice depends on the definition of standard management. If individuals who are eligible for NB32 would also engage in a weight loss programme when prescribed NB32 then the so-called intensive behaviour modification estimate might be more applicable. If this is not the case, then an estimate excluding intensive behaviour modification might be more appropriate. Of course, the estimate of 1.06 (0.84 to 1.33) is based on pooling both the trials with and without intensive behaviour modification and it is therefore tempting to infer that this represents clinical practice, where some do and some do not engage in weight loss programmes. This must be regarded with caution for a number of reasons, which include uncertainty as to the precise proportion who would engage in a weight loss programme and the degree of resemblance between such a programme and the intensive behaviour modification in COR-BMOD. Furthermore, costs of such intensive behaviour modification would also need to be considered in the economic model.

With regards to the economic model, one issue stood out: the structure and technical implementation of the company's model caused long run times (6 hours on average), and caused the model to crash on multiple computers. This hampered the company's and the ERG's ability to perform an appropriate PSA and the ERG's ability to check the model's validity and perform further scenario analyses (other than those that were described below). It should be considered whether simpler approaches (e.g. an individual-level state transition model) would have been more appropriate to reflect this decision problem, given the gain in transparency and that it would have been possible to reflect the condition-specific events in such a model. An individual-level state transition approach would potentially resolve most of the validity issues (e.g. the fact that BMI was not accurately reflected at each time period).

Apart from this, the ERG identified numerous issues of which the most important ones are summarised in Table 5.20. The ERG was able to adjust/correct some of these issues in its base-case. The ERG base-case ICERs (deterministic) of NB32 compared with standard management and orlistat ranged between £9,813-£10,510 and £38,871-£45,694 per QALY gained respectively. Subgroup analyses performed conditional on the ERG base-case, indicated that the ICERs (deterministic) of NB32 compared with standard management and orlistat were £10,535 per QALY gained and dominated respectively for T2DM patients and £9,594 and £25,744 per QALY gained respectively for non-T2DM patients. However, it should be noted that several issues remained unexplored (some of which were expected to be non-conservative) and thus the results should be interpreted in this context (i.e. with extreme caution). The interpretation and validity of the results are particularly hampered given that the company's model did underestimate TTD, did not incorporate behaviour modification interventions, bariatric surgery and re-treatment nor accurately reflected patients' expected quality of life and costs associated with resource use. As discussed in Section 5.2.12, the fact that BMI development was not accurately reflected in the model (due to lack of an updating event or integration of the BMI function) could significantly bias the results in favour of NB32. The large variation around the ICERs when different random numbers and sampled patient profiles are used is of particular concern. In two different model runs of the ERG base-case, the ICER varied by as much as £7,000 per QALY gained. It is therefore the ERG's view that the company's model is of very limited value for the current decision problem and that results are to be interpreted with extreme caution.

In conclusion, given that the deterministic ERG base-case ICER of NB32 versus orlistat is estimated to range between £38,871 and £45,694 per QALY gained (based on different random numbers and different samples of patients), and the remaining issues/methodological flaws highlighted above, uncertainty around the cost effectiveness estimates of NB32 remains substantial.

7.2 Strengths and limitations of the assessment

The majority of searches for eligible studies in the CS were well documented and easily reproducible. Searches were carried out on a good range of databases and carried out in accordance with the NICE 2013 guide to the methods of technology appraisal Sections 5.2.2 and 5.2.4. The strategies utilised recognised study design filters. Supplementary searches of conference proceedings and organisational websites, and the checking of references lists were undertaken by the company in order to identify additional studies not retrieved by the main searches.

Four good quality large RCTs for NB32 and 16 comparator trials were included in the submission. Analyses were presented for all patients and people with and without T2DM, including a large number of sensitivity analyses.

The economic model structure is similar to the assessment by Ara et al.[58], which is a Health Technology Appraisal report (2012) comparing different pharmacological treatments for obesity.

The main weakness of the CS was the use of mITT populations for the NB32 trials. These data overestimate the benefits of NB32 over placebo or orlistat when compared to the true ITT data.

The validity issues highlighted by the ERG, the technical implementation of the model, as well as the assumptions regarding TTD, lack of reflection of behaviour modification interventions, bariatric surgery and re-treatment, and inaccurate reflection of BMI hamper the interpretation and therefore question the validity of the results.

Furthermore, the ERG considers the model as unfit for purpose, due to its extremely long run times, the fact that it crashes on many computers, and the inability to perform PSA.

7.3 Suggested research priorities

An ongoing randomised trial in the US will be available in up to six years' time to provide data concerning the effect of NB32 on the occurrence of major adverse cardiovascular events (MACE) in overweight and obese adults with cardiovascular disease.[31] Further research will also be needed to ascertain the role of NB32 in patients who are overweight with comorbidities and patients of Asian ethnicity. Long term weight loss and maintenance should be investigated and any additional benefits of NB32 over and above intensive behaviour management clarified.

8. REFERENCES

[1] Orexigen. *Naltrexone-bupropion (prolonged release) for managing overweight and obesity [ID:757]: Submission to National Institute of Health and Clinical Excellence. Single technology appraisal (STA)*, January 2017 [accessed 11.1.17]. 267p.

[2] Moody A, Neave A. *Health survey for England 2015: adult overweight and obesity [Internet]*: Health and Social Care Information Centre, 2016 [accessed 3.2.17] Available from: http://www.content.digital.nhs.uk/catalogue/PUB22610/HSE2015-Adult-obe.pdf

[3] Baker C, Bate A. *Obesity statistics: briefing paper 3336 [Internet]*: House of Commons Library, Feb 2016 [accessed 31.10.16] Available from: researchbriefings.files.parliament.uk/documents/SN03336/SN03336.pdf

[4] Butland B, Jebb S, Kopelman P, McPherson K, Thomas S, Mardell J, et al. *Foresight. Tackling obesities: future choices – project report [Internet]*: Government Office for Science, 2007 [accessed 11.10.16] Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/287937/07-1184x-tackling-obesities-future-choices-report.pdf.

[5] National Institute of Health and Care Excellence. *Obesity: identification, assessment and management. NICE clinical guideline 189 [Internet]*. London: NICE, 2014 [accessed 12.10.16] Available from: https://www.nice.org.uk/guidance/cg189/chapter/1-recommendations

[6] World Health Organisation (WHO). *Obesity and overweight: fact sheet [Internet]*, June 2016 [accessed 10.10.16] Available from: http://www.who.int/mediacentre/factsheets/fs311/en/

[7] National Institute for Health and Clinical Excellence. *Naltrexone-bupropion (prolonged release) for managing overweight and obesity: Final scope [Internet]*. London: NICE, October 2016 [accessed 31.10.16]. 4p. Available from: https://www.nice.org.uk/guidance/GID-TAG486/documents/final-scope

[8] National Institute for Health and Care Excellence. *Naltrexone-bupropion (prolonged release) for managing overweight and obesity [ID757]: clarification letter*. London: NICE, 2017. 10p.

[9] Orexigen. *Naltrexone-bupropion (prolonged release) for managing overweight and obesity [ID757]: response to request for clarification from the ERG*, 2017. 55p.

[10] European Medicines Agency (EMA). *Assessment report for Xenical (Procedure number: EMEA/H/C/154/A-20/0057) [Internet]*, 2012 [accessed 26.10.16] Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Assessment_Report_-_Variation/human/000154/WC500134995.pdf

[11] National Institute for Health and Clinical Excellence. *Naltrexone-bupropion (prolonged release) for managing overweight and obesity: Response to consultee and commentator comments on the draft scope [Internet]*. London: NICE, October 2016 [accessed 31.10.16]. 4p. Available from: https://www.nice.org.uk/guidance/GID-TAG486/documents/scope-consultation-comments-and-responses

[12] Public Health England. Measurement of obesity [Internet]. 2016 [accessed 13.10.16]. Available from: http://www.noo.org.uk/NOO_about_obesity/measurement

[13] Canadian Agency for Drugs and Technologies in Health. *CADTH peer review checklist for search strategies [Internet]*. Ottawa: CADTH, 2013 [accessed 17.7.13]. 3p. Available from: http://www.cadth.ca/en/resources/finding-evidence-is

[14] National Institute for Health and Care Excellence. *Single Technology Appraisal: specification for manufacturer/sponsor submission of evidence [Internet]*. London: NICE, 2012 [accessed 31.1.17]. 76p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/Specification-for-manufacturer-sponsor-submission-of-evidence-June-2012.doc

[15] Scottish Intercollegiate Guidelines Network (SIGN). Search Filters [Internet]. Edinburgh: Healthcare Improvement Scotland, 2015 [accessed 19.1.17]. Available from: http://www.sign.ac.uk/methodology/filters.html

[16] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[17] R Core Team. The R project for statistical computing [Internet]. 2016 [accessed 22.11.16]. Available from: https://www.R-project.org/

[18] Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36(3):1-48.

[19] Haslam D. Obesity: current and emerging drug treatments. *Pharm J* 2015;294(7866):20068647.

[20] DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88.

[21] Lunn D, Thomas A, Best N, Spiegelhalter D. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000;10(4):325-37.

[22] Brooks S, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;7(4):434-55.

[23] Woodward P. *Bayesian analysis made simple: an excel GUI for WinBUGS.* Boca Raton, FL: CRC Press, 2011.

[24] Dias S, Welton N, Sutton A, Ades A. *NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials [Internet]*, 2011 (Updated: September 2016) [accessed 30.9.16] Available from: http://www.nicedsu.org.uk/TSD2%20General%20meta%20analysis%20corrected%202Sep2016v2.pdf

[25] Greenway FL, Fujioka K, Plodkowski RA, Mudaliar S, Guttadauria M, Erickson J, et al. Effect of naltrexone plus bupropion on weight loss in overweight and obese adults (COR-I): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* 2010;376(9741):595-605.

[26] Wadden TA, Foreyt JP, Foster GD, Hill JO, Klein S, O'Neil PM, et al. Weight loss with naltrexone SR/bupropion SR combination therapy as an adjunct to behavior modification: the COR-BMOD trial. *Obesity (Silver Spring)* 2011;19(1):110-20.

[27] Apovian CM, Aronne L, Rubino D, Still C, Wyatt H, Burns C, et al. A randomized, phase 3 trial of naltrexone SR/bupropion SR on weight and obesity-related risk factors (COR-II). *Obesity (Silver Spring)* 2013;21(5):935-43.

[28] Hollander P, Gupta AK, Plodkowski R, Greenway F, Bays H, Burns C, et al. Effects of naltrexone sustained-release/bupropion sustained-release combination therapy on body weight and glycemic parameters in overweight and obese patients with type 2 diabetes. *Diabetes Care* 2013;36(12):4022-9.

[29] Nissen SE, Wolski KE, Prcela L, Wadden T, Buse JB, Bakris G, et al. Effect of naltrexone-bupropion on major adverse cardiovascular events in overweight and obese patients with cardiovascular risk factors: a randomized clinical trial. *JAMA* 2016;315(10):990-1004.

[30] Halseth A, Shan K, Walsh B, Gilder K, Fujioka K. Method-of-use study of naltrexone sustained release (SR)/bupropion SR on body weight in individuals with obesity (In Press) [Data on file provided with the company's submission]. 2016.

[31] Orexigen. A multicenter, randomized, double-blind, placebo-controlled study assessing the occurrence of major adverse cardiovascular events (MACE) in overweight and obese subjects with cardiovascular risk factors receiving naltrexone SR/bupropion SR [Data on file provided with the company's submission]. 11 May 2015

[32] Orexigen. *Preliminary synopsis: single-dose pharmacokinetic study in hepatic impairment: a phase 1, open-label, parallel study to evaluate the pharmacokinetics of a single oral dose of extended-release combination of naltrexone and bupropion in subjects with normal hepatic function or varying degrees of impaired hepatic function (Naltrexone HCl / Bupropion HCl prolonged release) [PDF provided with the company's response to clarification]*, 2016

[33] Orexigen. *Preliminary synopsis: single-dose pharmacokinetic study in renal impairment: a phase 1, open-label, parallel study to evaluate the pharmacokinetics of a single oral dose of extended-release combination of naltrexone and bupropion in subjects with normal renal function or varying degrees of impaired renal function (Naltrexone HCl / Bupropion HCl prolonged release) [PDF provided with the company's response to clarification]*, 2016

[34] Orexigen. *A multicenter, randomized, double blind, placebo controlled study comparing the safety and efficacy of two doses of naltrexone sustained release (SR)/bupropion sustained release (SR) and placebo in obese subjects. (Clinical Study Report: NB-301) [PDF provided with the company's submission]*, 2010

[35] Orexigen. *A multicenter, randomized, double-blind, placebo-controlled study comparing the safety and efficacy of naltrexone sustained release (SR)/bupropion sustained release (SR) and placebo in subjects with obesity participating in a behavior modification program (Clinical Study Report: NB-302) [PDF provided with the company's submission]*, 2010

[36] Orexigen. *A multicenter, randomized, double-blind, placebo-controlled study comparing the safety and efficacy of naltrexone sustained release (SR)/bupropion sustained release (SR) and placebo*

in obese subjects (Clinical Study Report: NB-303) [PDF provided with the company's submission], 2010

[37] Orexigen. *A multicenter, randomized, double-blind, placebo-controlled study comparing the safety and efficacy of naltrexone 32 mg sustained release/bupropion 360 mg sustained release and placebo in obese subjects with type 2 diabetes mellitus (Clinical Study Report: NB-304) [PDF provided with the company's submission]*, 2009

[38] Orexigen. *Naltrexone-bupropion (prolonged release) for managing overweight and obesity [ID:757]: Appendices to accompany company evidence submission. Single technology appraisal (STA)*, January 2017 [accessed 11.1.17]. 268p.

[39] Roche. *Summary of product characteristics (SPC): Xenical 120 mg hard capsules [Internet]*, 2009 [accessed 26.7.16] Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000154/WC500058428.pdf

[40] Orexigen. Summary of product characteristics (SPC): Mysimba (Naltrexone/bupropion) [PDF provided with the company's submission]. 2016.

[41] Astrup A, Carraro R, Finer N, Harper A, Kunesova M, Lean ME, 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(6):843-54.

[42] Bakris G, Calhoun D, Egan B, Hellmann C, Dolker M, Kingma I. Orlistat improves blood pressure control in obese subjects with treated but inadequately controlled hypertension. *J Hypertens* 2002;20(11):2257-67.

[43] 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.

[44] Broom I, Wilding J, Stott P, Myers N. Randomised trial of the effect of orlistat on body weight and cardiovascular disease risk profile in obese patients: UK Multimorbidity Study. *Int J Clin Pract* 2002;56(7):494-9.

[45] Derosa G, Mugellini A, Ciccarelli L, Fogari R. Randomized, double-blind, placebo-controlled comparison of the action of orlistat, fluvastatin, or both an anthropometric measurements, blood pressure, and lipid profile in obese patients with hypercholesterolemia prescribed a standardized diet. *Clin Ther* 2003;25(4):1107-22.

[46] Derosa G, Maffioli P, Salvadeo SA, Ferrari I, Gravina A, Mereu R, et al. Comparison of orlistat treatment and placebo in obese type 2 diabetic patients. *Expert Opin Pharmacother* 2010;11(12):1971-82.

[47] Gotfredsen A, Westergren Hendel H, Andersen T. Influence of orlistat on bone turnover and body composition. *Int J Obes Relat Metab Disord* 2001;25(8):1154-60.

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

[49] Kelley DE, Bray GA, Pi-Sunyer FX, Klein S, Hill J, Miles J, 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.

[50] Lindgarde F. The effect of orlistat on body weight and coronary heart disease risk profile in obese patients: the Swedish Multimorbidity Study. *J Intern Med* 2000;248(3):245-54.

[51] Lucas CP, Boldrin MN, Reaven GM. Effect of orlistat added to diet (30% of calories from fat) on plasma lipids, glucose, and insulin in obese patients with hypercholesterolemia. *Am J Cardiol* 2003;91(8):961-4.

[52] Mathus-Vliegen EM, van Ierland-van Leeuwen ML, Bennink RJ. Influences of fat restriction and lipase inhibition on gastric emptying in obesity. *Int J Obes (Lond)* 2006;30(8):1203-10.

[53] Miles JM, Leiter L, Hollander P, Wadden T, Anderson JW, Doyle M, et al. Effect of orlistat in overweight and obese patients with type 2 diabetes treated with metformin. *Diabetes Care* 2002;25(7):1123-8.

[54] Reaven G, Segal K, Hauptman J, Boldrin M, Lucas C. 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.

[55] Swinburn BA, Carey D, Hills AP, Hooper M, Marks S, Proietto J, et al. Effect of orlistat on cardiovascular disease risk in obese adults. *Diabetes Obes Metab* 2005;7(3):254-62.

[56] Torgerson JS, Hauptman J, Boldrin MN, Sjostrom L. XENical in the prevention of diabetes in obese subjects (XENDOS) study: a randomized study of orlistat as an adjunct to lifestyle changes for the prevention of type 2 diabetes in obese patients. *Diabetes Care* 2004;27(1):155-61.

[57] National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal 2013 [Internet]*. London: NICE, 2013 [accessed 31.1.17]. 93p. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword

[58] 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.

[59] Davies MJ, Chubb BD, Smith IC, Valentine WJ. Cost-utility analysis of liraglutide compared with sulphonylurea or sitagliptin, all as add-on to metformin monotherapy in Type 2 diabetes mellitus. *Diabet Med* 2012;29(3):313-20.

[60] Burch J, McKenna C, Palmer S, Norman G, Glanville J, Sculpher M, et al. Rimonabant for the treatment of overweight and obese people. *Health Technol Assess* 2009;13 Suppl 3:13-22.

[61] Beaudet A, Palmer JL, Timlin L, Wilson B, Bruhn D, Boye KS, et al. Cost-utility of exenatide once weekly compared with insulin glargine in patients with type 2 diabetes in the UK. *J Med Econ* 2011;14(3):357-66.

[62] Copley V. *User guide: weight management economic assessment tool version 2 [Internet]*. Oxford: Public Health England, 2016 [accessed 11.8.16] Available from: http://www.noo.org.uk/visualisation/economic_assessment_tool

[63] Caro JJ. Discretely Integrated Condition Event (DICE) simulation for pharmacoeconomics. *PharmacoEcon* 2016;34(7):665-72.

[64] Orexigen. Clinical validation meeting for mysimba® NICE submission [Data on file provided with the company's submission]. 29 September 2016.

[65] Dare S, Mackay DF, Pell JP. Relationship between smoking and obesity: a cross-sectional study of 499,504 middle-aged adults in the UK general population. *PLoS One* 2015;10(4):e0123579.

[66] Gu Q, Burt VL, Dillon CF, Yoon S. Trends in antihypertensive medication use and blood pressure control among United States adults with hypertension: the National Health And Nutrition Examination Survey, 2001 to 2010. *Circulation* 2012;126(17):2105-14.

[67] Health and Social Care Information Centre. *Health survey for England 2013 (NS) [Internet]*, 2014 [accessed 2.3.17] Available from: http://content.digital.nhs.uk/catalogue/PUB16076

[68] National Institute of Health and Care Excellence. *Weight management: lifestyle services for overweight or obese adults. NICE public health guideline PH53 [Internet]*. London: NICE, 2014 [accessed 12.10.16] Available from: https://www.nice.org.uk/guidance/ph53

[69] Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. *Value Health* 2011;14(4):539-45.

[70] MIMS. Orlistat [Internet]. 2016 [accessed 21.11.16]. Available from: http://www.mims.co.uk/drugs/nutrition/obesity/orlistat

[71] Curtis L, Burns A. *Unit Costs of Health and Social Care: 2015 [Internet]*. Canterbury: Personal Social Services Research Unit (PSSRU), 2015 [accessed 18.11.16] Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2015/index.php

[72] Department of Health. *NHS reference costs 2014-2015 [Internet]*. London: Department of Health, 2015 [accessed 22.9.16] Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015

[73] Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. *Health Technol Assess* 2009;13(34):1-74, 75-118.

[74] Robinson M, Palmer S, Sculpher M, Philips Z, Ginnelly L, Bowens A, et al. Cost-effectiveness of alternative strategies for the initial medical management of non-ST elevation acute coronary syndrome: systematic review and decision-analytical modelling. *Health Technol Assess* 2005;9(27):iii-iv, ix-xi, 1-158.

[75] Youman P, Wilson K, Harraf F, Kalra L. The economic burden of stroke in the United Kingdom. *PharmacoEcon* 2003;21 Suppl 1:43-50.

[76] Diabetes UK. Cost of diabetes [Internet]. [accessed 8.11.16]. Available from: http://www.diabetes.co.uk/cost-of-diabetes.html

[77] Diabetes UK. Number of people with diabetes reaches over 4 million [Internet]. 2016 [accessed 31.10.16]. Available from: https://www.diabetes.org.uk/About_us/News/Number-of-people-with-diabetes-reaches-over-4-million/

[78] World Health Organisation (WHO). Cardiovascular diseases (CVDs): fact sheet [Internet]. 2016 [accessed 11.10.16]. Available from: http://www.who.int/mediacentre/factsheets/fs317/en/

[79] World Health Organisation (WHO). Types of cardiovascular disease [Internet]. 2004 [accessed 22.11.16]. Available from: http://www.who.int/cardiovascular_diseases/en/cvd_atlas_01_types.pdf

[80] Joint Formulary Committee. *British National Formulary: Adverse reactions to drugs [Internet]* London: BMJ Group and Pharmaceutical Press, 2016 [accessed 22.9.16]. Available from: http://www.evidence.nhs.uk/formulary/bnf/current/guidance-on-prescribing/adverse-reactions-to-drugs

[81] Davis S, Stevenson M, Tappenden P, Wailoo AJ. *NICE DSU technical support document 15: cost-effectiveness modelling using patient-level simulation [Internet]*, 2014 [accessed 7.3.17] Available from: http://www.nicedsu.org.uk

[82] Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force 7. *Value Health* 2012;15(6):843-50.

[83] Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Moller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force 4. *Value Health* 2012;15(6):821-7.

[84] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

APPENDIX 1: DETAILS OF ERG ANALYSES (FOR VALIDATION PURPOSES)

Adjusted cells are printed in *Italics*

Fixing errors

1. Fixing errors consisted of using a weight regain period of 1.5 years after which weight is instantly regained, to reflect the three year linear weight regain assumption.
   Efficacy I 117

Fixing violations

2. Using the ITT data instead of mITT data; and based on the COR-I and COR-DM trials only
   Efficacy F35:M36, I53:56, I74, I88:92, AS6:BI10017
3. Using a relative risk instead of mean difference to extrapolate the difference between treatments in change from baseline weight from the secondary to the primary assessment.
   Efficacy I53:56
4. The natural history model to predict BMI is calibrated to reflect the baseline BMI distribution as observed in the COR trial programme.
   DICE equations W108, G118
5. Adjusting the baseline age (dependent on T2DM status), proportions of females (dependent on T2DM status), proportion of smokers, proportion receiving statins (dependent on T2DM status), proportion receiving anti-hypertensive medication (dependent on T2DM status) and proportion receiving aspirin.
   Controls J27:K42, DICE equations AX15, BA15
6. Removal of GP visit for standard management.
   Non-drug costs J90

Matters of judgment

7. Assuming weight regain towards baseline BMI instead of predicted BMI.
   DICE equations I 415, G:J415, G:J417, D419, D486, I486, D553, I553, D614, I614, D682, I682
8. Removing linear scaling assumption of TTD for orlistat.
   Treatment duration AO7:AP221,
   Range named td_first_response_lookup_int_b

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

Pro-forma Response

ERG report

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

You are asked to check the ERG report from Kleijnen Systematic Reviews Ltd to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm, 17 March 2017** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Standard Management

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 11 of the ERG report and similar sentences throughout: *"The company assumed that standard management was more like that in the COR-I and COR-II trials in which patients received advice on diet and exercise but were not allowed to participate in a weight loss programme."* Page 12 of the ERG report and similar sentences throughout: *"No trials were identified that compared NB32 directly with orlistat or with different types of behavioural interventions."* Page 13 of the ERG report and similar sentences throughout: *"A comparison between NB32 (plus standard management) versus intensive behaviour modification is missing."* Page 13 of the ERG report and similar sentences throughout:	It should be stated that standard management for the treatment of obesity in the UK should be considered as comparable to that described in the COR I and COR II clinical trials. We request that the ERG acknowledge that standard management encapsulates several dietary and behavioural interventions in clinical management of varying intensity, and that this is the definition adopted in the submission, rather than assumptions of one control arm better representing standard management than another. As part of this, we also request that the ERG reconsider their conclusion that the control arms of any one trial, particularly the COR-I and COR-II trials, are not reflective of standard management in clinical practice and do not therefore provide comparative data for NB32 plus standard management versus standard management alone (of varying intensity depending on trial).	Clinical validation has confirmed that the diet and exercise (non-drug) treatment currently received by patients alongside pharmacological treatment varies by postcode, and can be delivered by dietician, GP or WeightWatchers®, but that the diet and exercise treatment in the COR-I and COR-II studies was a good reflection of the average. Patients in the COR I and II trials undergoing standard management, 'were instructed to follow a hypocaloric diet representing a deficit of 500 kcal per day based on the World Health Organization algorithm for calculating resting metabolic rate. Adjusted body weight was used to calculate energy needs because subjects were 120% greater than ideal body weight. Subjects received written instructions on behavioural modification techniques. Patients were encouraged to increase physical activity, with a prescription for walking starting with at least 10 minutes on most days of the week, and increasing this gradually to 30 minutes on most days of the week throughout the study. They were	Not a factual error. The assertion that the diet and exercise treatment in COR-I and COR-II was a good reflection of standard management can be challenged, at least partly, according to the similarity between the company's own description of one option of what is currently received, i.e.WeightWatchers®, and the intense behaviour modification not available in COR-I or COR-II, which includes: "…group meetings (10 to 20 patients per session) lasting 90 minutes (including weigh-in)…". It is on this basis i.e. the exclusion of such weight loss programmes from the COR-I and COR-II trials that the ERG questions the appropriateness of these trials and suggests that COR-BMOD might be more appropriate to represent standard management.

"Standard management in the UK might be better reflected by COR-BMOD." Page 29 of the ERG report and similar sentences throughout: '*In the COR-I and COR-II trials participants were not permitted to engage in weight loss programmes other than the prescribed programme of diet modification and exercise advice. This represents a more minimal approach to standard management than might be expected in practice. The COR-BMOD trial could be seen as best practice for standard management in that a more intensive intervention was delivered'*		encouraged to lose weight and maintain weight loss, and were encouraged to follow the prescribed programme'. Furthermore, when asked about the more intensive diet and exercise intervention used in the COR-BMOD study, a clinician confirmed that this is similar to the type of non-drug treatment used for Tier 3 patients in NHS Practice. Intense behavioural modification within the BMOD study is described as 'all patients were to participate in an intensive behaviour modification program that included three components: dietary instruction, closed group sessions, and prescribed exercise. Behaviour modification consisted of group meetings (10 to 20 patients per session) lasting 90 minutes (including weigh-in) weekly for the first 16 weeks, every other week for the next 12 weeks and monthly thereafter for up to 28 sessions. They included instructions to consume a balanced deficit diet and to increase to 180 min/week of planned, moderately vigorous, physical activity. Dietary instructions were provided at baseline (Day 1).' NICE Clinical Guideline 89 (Obesity: Identification, assessment and	

		management) defines Tier 3 obesity as a BMI of 35 to 39.9 kg/m^2 and in NICE Public Health Guideline 53 (Weight management: lifestyle services for overweight or obese adults) Tier 3 services are defined as specialist weight management. As such, we consider that all trials include a standard management arm which is a fair reflection of that seen in clinical practice within NHS England.	

Issue 2 The ERG's misinterpretation of the preferred analysis to inform the indirect treatment comparison

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 18 of the report, it states: *"The ERG considers that the use of the ITT population (instead of the mITT) to inform treatment response and weight loss would have been both more appropriate and more conservative."* The ERG's preference for the ITT population is stated	We propose the ERG reconsider their preferred population for use within the ITC, particularly noting the differences in preferred population for the ITC as an independent analysis of the *de novo* model and as an analysis to specifically inform the *de novo* model.	The ITC was undertaken to inform the relative efficacy measures required in the *de novo* model. The decision to undertake the analysis in this manner was pragmatic, given the time taken to produce such analyses is lengthy. An ITC to demonstrate the relative efficacy of treatments in the absence of a co-existing *de novo* economic model clearly should utilise the ITT population, as is standard in clinical evidence synthesis guidelines. In this respect, we agree with the ERG that the ITT population is preferred when considering the	This is not a factual inaccuracy. The ERG preferred to use the ITT population for the ITC (see section 4.5.1 for more details). See Table 4.33 in the ERG report for a summary of the ITC estimates used in the ERG preferred analyses. No amendments to the ERG report are needed.

throughout the report. However, the ERG do not distinguish between their preferred population in relation to: • the indirect treatment comparison (ITC) independent of the *de novo* model, and • the ITC for use within the *de novo* model. The use of the mITT population for the ITC was pragmatically chosen, given its role in the derivation of relative efficacy measures in the *de novo* model. The ERG's interpretation of which population is preferred is unclear and misleading.		ITC as a "stand-alone" analysis. The ITC presented in this appraisal was undertaken specifically to inform the *de novo* model. Within the model, TTD data from the ITT population are used to capture the proportion of patients who discontinue ahead of primary assessment. Clearly, the majority of patients who do not have a recorded weight loss at primary assessment (or beforehand other than baseline) are likely to have discontinued, and are therefore captured within the TTD curve. If the ITT analysis were utilised to inform the relative treatment effects in the *de novo* model, the proportion of non-responders would be over-estimated – patients who had previously discontinued would be erroneously double-counted in this estimation of response. Clarification of the use of the mITT population in the ITC is important for committee understanding – i.e. that the mITT population is most appropriate for establishing outcomes in the *de novo* model without double-counting non-responders.	
On page 81 of the report, the ERG used the ITT population with baseline-carried forward analysis (BCFA). The use of BCFA is likely to bias results against NB32. No reason for the preference of BCFA has been provided.	We request that the ERG reconsider their chosen method of data imputation in their preferred analyses, in light of the clarification presented to highlight their misinterpretation of the BCFA. If the ERG still consider that BCFA is appropriate, we ask that the	In the COR-I CSR the BCFA analysis is defined in the following way: *"An additional sensitivity analysis of the percent change from baseline of total body weight was conducted in which endpoint was defined as Week 56 measurement, and endpoint was defined as the baseline for subjects who discontinued active study drug*	Not a factual error. The choice between the two methods of data imputation was mainly because some data were not available for ITT-Imp (= all randomised patients with weight regain imputation method analysis), whereas for the BCFA method all data were available (see

	ERG provide justification on why it is their preferred approach.	*prior to Week 56 (i.e., the percent change from baseline was equal to zero for these subjects). This analysis is referred to as the baseline-carried-forward-analysis, and was conducted on all randomized subjects."* Hence in this analysis it is only completers that experience benefit from treatment; and that any patient who discontinues treatment prior to week 56 would receive no treatment effect. This is evident in the COR-I study where the number of responders on NB32 was 226 patients in our analysis but falls to 180 patients in the ITT-BCFA, despite a wider group of patients being considered. We urge the ERG to acknowledge that this is biased against NB32 as patients who discontinued prior to week 52 in the orlistat studies would have their last observation carried forward (LOCF), <u>not</u> their baseline observation. We understand that the use of LOCF is not a perfect approach to measure weight loss at 56 weeks. However, it is more reflective of actual weight loss and, crucially, it also allows the method of imputation to be consistent with the orlistat studies. The use of LOCF is standard in the reporting of outcomes in obesity studies, and was also used by the majority of studies in the mixed-treatment comparison performed by Ara *et al.*	Table 4.27 of the ERG report).

Issue 3 The ERG's unjustified and incorrect implications of bias in favour of NB32

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 17 of the report, it states: *"The ERG considered the BMI sampled in the model and compared it with the baseline BMI in the COR trial programme and concluded that baseline BMI is vastly underestimated in the economic model. This is also reflected in the average baseline weight of 92kg in the model, while the averages ranged between 99kg and 105kg in the COR trial programme. Given that BMI is included as a predictive factor for utility, T2DM, cardiovascular events and death, the utility values and the time to these events in the model are overestimated, likely inducing bias in favour of NB32."* This is a clear and illustrative case of the ERG identifying a limitation in the model and using this to imply that the model is biased in favour of NB32, without clear thought, evidence or explanation.	We propose the ERG amend their statement to the following (introduced wording underlined): *"The ERG considered the BMI sampled in the model and compared it with the baseline BMI in the COR trial programme and concluded that baseline BMI is underestimated in the economic model. This is also reflected in the average baseline weight of 92kg in the model, while the averages ranged between 99kg and 105kg <u>across trials in</u> the COR trial programme. Given that BMI is included as a predictive factor for utility, T2DM, cardiovascular events and death, the utility values and the time to these events in the model are overestimated. As the ERG changed baseline BMI in the model to reflect baseline COR trial BMI for their preferred analysis, it was possible to assess the implications of baseline BMI values for model results"*	We feel the ERG should demonstrate that a balanced and informed approach to consider the implications of model assumptions has been taken, if they are to make claims about bias in their report. We invite the ERG to consider the following in their appraisal of the implications of baseline BMI: • Weight change is measured as % change from baseline • Higher baseline BMI implies shorter times to clinically relevant events (such as onset of T2DM). Further, as the ERG changed baseline BMI in the model to reflect baseline COR trial BMI for their preferred analysis, it should be perfectly possible and within the ERG's remit to assess the implications of baseline BMI values for model results.	This is not a factual inaccuracy. The ERG attempts to accompany their criticism of any issues by a (likely) direction of possible bias. In this specific case, given the limitations of the model (see ERG report section 5.2.12) it was not feasible to show the implications of all these issues by conducting model analyses. Hence, the ERG provided the "likely" direction of bias based on its understanding of the economic model. Therefore, the ERG added "likely" when describing the direction of bias. Given that the company did not provide any analyses to show that the "likely" directions of bias, as described by the ERG, are incorrect, no amendments to the ERG report are needed.
On page 110 of the report, it states:	The ERG should amend table 5.20 to clarify that the pooling of all the COR studies (including COR-	The inclusion of the COR-BMOD study in the ITC would result in an over estimate of the absolute number of non-T2DM	Not a factual error.

"it was inappropriate to pool the proportion of responders to NB32 treatment from all COR studies, including BMOD. By doing so, the proportion of responders to NB32 is over-estimated. This is supported by response rates for treatment with NB32 versus placebo as presented in Table 4.8. As a result, it is the ERG's view that response rates to NB32 are likely to be over-estimated as a consequence of the pooling method." A similar statement is made on page 112: *"it is the ERG's view that mean change in weight for patients treated with NB32 is likely to be over-estimated as a consequence of the pooling method (Table 4.8)."* In table 5.20 it is also stated that the pooling of studies (COR-I, COR-II, COR-BMOD) is likely to bias results in favour of NB32. Though differences are noted between the COR trials, the ERG's statement is factually inaccurate. The difference in relative effect estimates observed in the BMOD-study are much more conservative for the comparison of NB32	BMOD) was a <u>conservative assumption</u>. The ERG should also amend text in sections 5.2.6 iii) and iv) to consider the difference in the <u>relative effects</u> observed in the COR-BMOD study which result in a more conservative <u>relative effect</u> estimate between NB32 and the two comparators.	responders for NB32 – this increase would also be seen for orlistat when the relative effect from the ITC is applied. Furthermore, the relative effect of NB32 compared to standard management in the COR-BMOD study favours NB32 less than the COR-I and COR-II studies. The base-case analysis performed in the ITC therefore provides a more conservative estimate of the relative effects between NB32 and orlistat than SA3 where treatments with intensive BMOD were excluded (table 4.23, page 73 and table 4.25, page 75). This is also seen in the ERG's preferred analysis (table 4.33, page 81) where COR-I was the only study from the COR trial programme used to estimate the relative effects of NB32 against the two comparators.	

against standard management (table 4.8, page 46), which subsequently also produces more conservative relative effect estimates in the ITC when NB32 (pooled across the three studies rather than, say, just COR-I) is compared to orlistat in non-T2DM patients.			

Issue 4 The ERG's misinterpretation of model assumptions for time to treatment discontinuation

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 110 The ERG identified three publications reporting TTD for orlistat, which stated that orlistat TTD was longer than the 12.29 months estimated by the model, with the studies reporting that the proportion of patients still receiving orlistat at 12 months was >50%. This does not necessarily represent what happens in clinical practice. On page 110 of the report, it states: *"For the derivation of the orlistat TTD, the KM estimates for NB32 TTD for the first 16 weeks were linearly scaled to fit the first 12 weeks of orlistat treatment. This was justified by the different time*	The ERG should include within their interpretation of the TTD for orlistat that any uncertainty is bi-lateral i.e. the TTD for orlistat may be under- or over-estimated.	The studies identified by the ERG are relatively small investigator led trials that, based on expert clinical opinion, do not reflect standard practice in the UK. In the UK, many patients are unable to adhere to the low fat diet required for orlistat treatment and as a result stop taking orlistat due to experiencing adverse events preventing continuation of treatment. The ERG states that they consider TTD to be under-estimated for orlistat, as a direct result of the linear scaling approach. However, the known toxicity profile for orlistat (as highlighted by consultation	This is not a factual inaccuracy. No amendments to the ERG report are needed. See section 5.2.6 ii) of the ERG report for more details.

to primary assessment, and the fact that for NB32, the first four weeks include a titration period. The ERG believes that this linear scaling may further under-estimate orlistat TTD, resulting in worse effectiveness (patients will stop losing weight and start weight regain sooner), but also in decreased costs associated with orlistat and the effect of this is therefore ambiguous." The ERG fail to acknowledge that the implementation of TTD for orlistat may under- or over-estimate costs and associated effects. Hence, this statement is misleading.		with Prof. Wilding) suggests that patients in practice may discontinue treatment with orlistat at a greater rate than NB32 – particularly in the first weeks of treatment. Consequently, TTD for orlistat may be over-estimated yet this is not acknowledged by the ERG. This amendment to the ERG report is important to highlight to the committee the true nature of the uncertainty surrounding TTD for orlistat.	
On page 109 of the report, it states: *"It is the ERG's view that TTD for NB32 and orlistat may be under-estimated. The ERG wishes to highlight that the under-estimation of TTD leads to an under-estimation of costs."* An under-estimation of TTD leads to an under-estimation of costs for both treatment arms – the statement from the ERG should be amended to include this, as it is otherwise misleading.	The ERG should amend their statement for clarity to the following: *"It is the ERG's view that TTD for NB32 and orlistat may be under-estimated. The ERG wishes to highlight that the under-estimation of TTD leads to an under-estimation of costs for both NB32 and orlistat."*	Minor amendment of the ERG's statement is requested to aid clarity and understanding.	This is not a factual inaccuracy. No amendments to the ERG report are needed.
On page 109 of the report, it states: *"The end of the NB-CVOT study was used as the maximum TTD, whether patients in that study had discontinued*	It is clear that the implementation of TTD within the model is based on the availability of evidence, and that modelling treatment discontinuation up until the end of	Our implementation of TTD is consistent with the gold standard study conducted by Ara *et al.,* wherein it is stated: *"In the base case we assume that all*	This is not a factual inaccuracy. No amendments to the ERG report are needed.

or not." No further explanation of why this modelling assumption was applied is presented here. This is misleading, as rationale for why this modelling assumption was made has previously been communicated to the ERG via clarification question response and within the CS.	evidence is consistent with previous studies. As such, the reasons for which the model incorporated treatment discontinuation at the end of available evidence should be incorporated into the ERG's critique.	*patients are withdrawn from active treatment at 12 months as this is the end point for our evidence."* The end of the NB-CVOT study constitutes the end of our evidence, hence we adopted the same approach as Ara *et al.* in our model. In the systematic review of previous studies undertaken by Ara *et al.*, it was stated: *"The duration of treatment modelled* [in the 16 identified studies] *was generally 1 year, although one study used a lifetime of treatment, and one used 6 months of orlistat weight loss followed by a 6-month maintenance period."* As described in the CS, patient-level data from the NB-CVOT study comprise the best available evidence for NB32 treatment duration assumptions beyond 1 year in clinical practice. The modelling of TTD until the end of available evidence is reflective of many previous economic analyses undertaken in obesity.	
On page 109 of the report, it states: *"However, these TTD estimates* [identified in studies assessed by the ERG for orlistat] *were not conditional on response to treatment (primary and secondary response assessments) and therefore have to be interpreted with caution, but reported response rates in*	The ERG should clearly state the limitations in these studies, or otherwise remove consideration of these studies in stating their point.	The associated response rate for placebo patients in Bakris *et al.* is 22.6%, and in Broom *et al.* 24.3% - clearly higher than the associated rate for placebo patients in COR I and COR DM (16% and 18.9%, respectively). The relative effect of orlistat is therefore lessened within these studies.	This is not a factual inaccuracy. No amendments to the ERG report are needed.

two of these studies suggest that a significant proportion would still have continued treatment based on their response (45.7% response rate as measured by patients achieving >5% weight loss in Bakris et al.[42], 55.6% in Broom et al.[44]; Berne et al.[43] did not report response rates with the same level of weight loss).” The ERG fail to acknowledge the issues with these individual studies, which is misleading.		Both Broom *et al.* and Berne *et al.* included lead-in periods, which were not a feature of the COR trial programme. Regrettably, lead-in periods were a feature of historic trials of orlistat, but were not included in the COR trial programme as inclusion of lead-in periods is not reflective of clinical practice. These limitations of individual studies should be made clear by the ERG, as omission suggests much higher response rates for orlistat than those that would be seen in clinical practice.	

Issue 5 Description of model as “unfit for purpose” based solely on run time and ERG computing capacity

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 21 of the report, it states: *“Furthermore, the ERG considers the model as unfit for purpose, due to its extremely long run times, the fact that it crashes on many computers, and the inability to perform PSA.”* Similar statements feature throughout, including the following statement on the	We request that the ERG removes the statements highlighted, and any similar statements throughout the report. We suggest that the following text could be used to replace the highlighted text: *“The ERG considers model run time and computational power requirements as detrimental factors to their review of the company’s model, which led to difficulties in performing the usual number of PSA runs*	The ERG’s statements are both misleading and factually incorrect. For the ERG to report that there is “inability to perform PSA” is incorrect. The model allows the user to perform PSA as per NICE requirements. That there is computational burden associated with running large numbers of PSA iterations is different to “inability to perform PSA”, and the difference is important for the ERG’s audience. We were clear to explain and set out our rationale for the model type and structure in Section 5.2 of the CS. It is always preferable, for the manufacturer and the ERG, to use the	This is not a factual inaccuracy. No amendments to the ERG report are needed. See sections 5.2.10 and 5.2.12 for reasons why the ERG believes the model is unfit for purpose and the ERG was unable to perform PSA ; amongst others the following quotes: “The ERG recorded run times between 450 and 600 hours of a PSA with 100 individual randomly sampled patients and 1,000 PSA runs. These numbers of

same page: *"It is therefore the ERG's view that the company's model is of very limited value for the current decision problem and that results are to be interpreted with extreme caution."* These statements are factually inaccurate, unnecessarily critical of the model and grossly misleading. Accusing the *de novo* model of being *"unfit for purpose"* or of *"very limited value"* is highly inappropriate as the justification for this inflammatory language seems to be based solely on run time and computational power.	*within the time-frame allocated. However, we note that an individual-level approach to modelling is appropriate for the decision problem, and both the company's model and the ERG's suggestions for modelling are associated with run times and computational burden that are greater than cohort-level analyses, that are typically, but not always, appropriate for NICE STAs. The limitations the ERG faced were in terms of testing the full range of results sense-checks that are quick to run in a cohort state-transition model.* *The ERG could and did fully assess the model code and logic, and the suitability of the approach to address the decision problem. Only one instance of what could be described as an error was found. The model type and structure is based on and reflects a 2012 Health Technology Assessment report comparing different pharmacological treatments for obesity"*	simplest and most transparent model possible to adequately address the decision problem, and this is the approach that we have demonstrably taken. "Long run time" is a vague criticism, and if the ERG is to criticise the manufacturer for their experience of model execution, they should be careful to be both specific and clear in stating their reasoning. Typically, cohort models are appropriate for NICE STAs, and in such models, model execution for the cohort is comparable in burden to model execution for one individual in a standard approach to individual-level modelling. We set out the reasons why a cohort approach would have been unsuitable here, and the ERG seem to be of the same mind. The ERG's suggestion for an alternative "simpler" model to the company model (and, implicitly, the model built by Ara et al) is a discrete-time individual-level model. As covered under Issue 6, it is not at all clear that the ERG's proposed approach would provide quicker model run time. There is a potentially dangerous precedent set if treatments for diseases and conditions that require an individual-level model are penalised in the NICE STA process on the basis of the relative computational challenges of individual-level versus cohort-level modelling approaches, with phrases such as "unfit for purpose" used by ERGs to describe internally and externally valid models associated with necessary but mitigatable run time challenges.	sampled patients and PSA runs are still too low to obtain a reliable result. The ERG acknowledges that in any model study trade-offs are made between validity and reliability of the result and practical considerations. However, companies should provide a submission that is compliant to the NICE decision making process in which probabilistic model results are preferred, and the model is assessed by the ERG in a period of eight weeks. For this model, it was unfeasible for the ERG to perform an adequate assessment of the model's probabilistic results within the time frame of a NICE submission." "The ERG wishes to highlight that it considers the model submitted by the company as unfit for purpose. The implementation in DICE resulted in extremely slow runtimes (6 hours on average, but with occasional model run times of 10 hours, depending on computer specifications) for the deterministic analysis only. It should also be noted that the model crashed on most of the computers that it was tried on." Regarding the comment on precedent setting, the ERG wishes to stress that for some decision problems individual patient-level modelling is the preferred approach. The ERG agrees that for this decision problem individual patient-level modelling is appropriate. For this submission, the ERG criticised the implementation of the

			individual patient-level approach, not the individual patient-level approach itself.

Issue 6 Lack of justification for ERG's suggested alternative modelling approaches

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 19 of the report, it states: "*It should be considered whether simpler approaches (e.g. an individual-level state transition model) would have been more appropriate to reflect this decision problem, given the gain in transparency and that it would have been possible to reflect the condition-specific events in such a model. An individual-level state transition approach would potentially resolve most of the validity issues (e.g. the fact that BMI was not accurately reflected at each time period).*" On page 95 of the report, it states:	We propose that of the cited text, the only part suitable for *amended* inclusion report is the following: *"The ERG considered it reasonable to use the economic model by Ara et al., (comparing different pharmacological treatments for obesity) as a starting point <u>and key source</u> for the current analysis."*	The ERG's assessment of the suitability of the submitted model and proposal for a more viable alternative is vague, unjustified and misleading. Consideration of a few implications of the ERG's words illustrate this. <u>It is not clear if or how the ERG's suggested alternative is simpler</u> From the ERG's casual wording, it is not at all clear that their suggested "simpler" approach is in fact simpler. The ERG suggest an individual-level state-transition approach. It seems likely that rather than help what the ERG see as the key limitation of the analysis, the ERG's suggested alternative would in fact increase model run time. The number of calculations required per patient per run would be expected to increase due to the vast number of health states that would be required to specify a health state for each possible set of patient conditions. Is the ERG proposing different health states for BMI categories? This would be curious given the input data. How would the ERG categorise BMI and how would they justify their choices? Given the decision problem and availability of	This is not a factual inaccuracy. No amendments to the ERG report are needed. The sections quoted by the company, from the ERG report, in the far left column already provide the ERG's response: "It is unclear to the ERG why a DES approach is preferred over for instance an individual-level state transition model" Therefore, the ERG proposes that "It should be considered whether simpler approaches (e.g. an individual-level state transition model) would have been more appropriate to reflect this decision problem" The statement by the company (in the third column): "For the ERG to describe the submitted model as not transparent" is factually incorrect, the ERG does not

"It is unclear to the ERG why a DES approach is preferred over for instance an individual-level state transition model. However, the ERG considered it reasonable to use the economic model by Ara et al., (comparing different pharmacological treatments for obesity) as a starting point for the current analysis." These statements are misleading and factually inaccurate.		data, we were at pains to take the simplest approach possible to adequately address the decision problem. We have demonstrated our rationale and justifications in the CS. It is vital that the ERG employs a similar level of scrutiny if they are to claim that there is an alternative that is both adequate and simpler. <u>It is not clear if or how the ERG's suggested alternative would be more transparent</u> The ERG states that their alternative approach would bring a "*gain in transparency*". Though the ERG's suggestion for an alternative model is so vague, it is hard for us to give a detailed critique, this seems unlikely given the apparent complexity and compromise that would be required for a state-transition approach. For the ERG to describe the submitted model as not transparent is both misleading and disappointing. We were conscious to clearly and transparently structure the model, with minimal and clear use of background Visual Basic® code. This facilitated the ERG to check model logic, and the benefits of this good practice have manifested in the complete absence of logic errors identified in the model (**Error! Reference source not found.** describes our objection to the ERG's description of the one "error" they identified as an "error"). If the ERG truly feel the model is not transparent, a clear and fair justification for this claim is not presented in the report. Further, it is not clear or justified that the ERG's alternative approach would be in any way more transparent.	describe the model as not transparent; rather, it raises the question of whether alternatives would be more transparent.

		<u>It is not clear that the ERG have considered the validity implications of their suggested alternative approach</u> The ERG are critical of "*validity issues*" in the cited text and elsewhere in the report, and mainly focus their critique on the model's approach of updating conditions upon the incidences of events, instead of at regular time intervals; "*(e.g. the fact that BMI was not accurately reflected at each time period)"*. What the ERG do not say is that it would be perfectly possible to update conditions at regular time intervals within the submitted DES model structure. In this light, the entire sentence "*An individual-level state transition approach would potentially resolve most of the validity issues (e.g. the fact that BMI was not accurately reflected at each time period).*" is utterly misleading, and highly concerning. A consequence of updating conditions at regular intervals in the submitted model would be an increase in model run time. The ERG should consider whether they feel this would be justified. <u>The ERG must justify any criticisms and claims for a preferable alternative</u> We suggest that the ERG must explain what their preferred individual-level state transition model approach would comprise and how it would a) be simpler than our DES approach, given the available data and the requirement to reflect patient heterogeneity appropriately, and b) contribute to favourable model run times.	

		If the ERG cannot or does not do this, we expect the ERG to indicate to the committee that they endorse the company's model, while expressing their frustrations with the computational challenges they faced in their review.	

Issue 7 The one "error" that the ERG identified and the implications of the ERG's proposed "solution"

Description of problems	Description of proposed amendments	Justification for amendments	ERG Response
On page 17 of the report, it states: *"Furthermore, the linear weight regain over the time-span of three years was implemented incorrectly in the model in that the weight regain occurs instantaneously at the end of the three year period."* Elsewhere in the report, this "error" is also referred to, such as on page 95: *"Furthermore, the linear weight regain over the time-span of three years was implemented incorrectly in the model where, in fact, the weight regain occurs*	We ask that the ERG re-consider their correction of this "error" as we do not consider the original implementation to be an error, whereas we do consider the ERG's "solution" to introduce inconsistency. Further, we request that the ERG notes that their critique of model logic and code identified no errors. We request that this is noted in the summary section of the ERG report, and wherever the quality of the model is described.	The implementation of linear weight regain (over a three-year period) affects two features of the model; a) the derived utility for patients, and b) the predicted times to events. Below the impact on each of these is described, as the ERG have misinterpreted how weight regain was applied in the *de novo* model. For utility values, the QALYs gained following treatment discontinuation until regain are derived using an average between the utility at the beginning of the regain period and at the end of regain period (via linear interpolation). For example, a patient with a utility of 0.8 at the start of regain and 0.78 at the end of regain would accrue 3*(0.8+0.78)/2 = 2.37 undiscounted QALYs. This is shown as part of the equations to calculate QALYs in the *de novo* model: di_QALYs + ((di_time - di_prev_event_time) / di_year_length) * AVERAGE(di_previous_utility, di_utility)	This is not a factual inaccuracy. No amendments to the ERG report are needed. The ERG is concerned about the estimated times to subsequent events, which are hindered by the instantaneous application of weight regain. Therefore, the ERG prefers to maintain the 'fix' (which is not ideal), in the ERG base-case. Please, note that a scenario applying the 3 year weight regain period is provided as exploratory analysis (ICER difference is less than £1,000). Other errors have been identified by the ERG, for example a discrepancy in the reporting of the odds ratio of orlistat vs NB32 proportion of responders that

instantaneously at the end of the three year period. The ERG incorporated adjustments in its base-case to reflect a linear weight regain over three year." The "error" is stated on page 139, where fixing an error is defined as: *"correcting the model were the company's submitted model was unequivocally wrong"* To describe the company's model logic here as an "error" is incorrect. Further, the ERG's proposed solution introduces error.		Thus, the instantaneous application of regain does not directly hinder utility (and hence QALY) derivation. For the times to subsequent events, the previous time is adjusted according to the updated covariate values (i.e. BMI) at regain. This is appropriate, and takes into account the regain period logically. The ERG's proposed solution (that is, changing the regain time to 1.5 years) introduces unintended inconsistency. The implementation suggests that weight is fully regained at 1.5 years, and that the predicted times to next events are calculated from 1.5 years <u>not</u> 3 years as per the base-case assumption. The ERG's "solution" causes underprediction of utility across all treatment arms, as the time period over which patients are regaining weight is halved. Regarding time to event, the ERG's "correction" causes the predicted times to events to be underpredicted. We ask that the ERG re-consider their correction of this "error" as we do not consider the original implementation to be an error, whereas we do consider the proposed solution to introduce further error. Reversion of the base-case model setting to a three-year weight regain is expected to result in minimal changes to model results, but clarification should aid committee understanding of why the original implementation should not be considered as an error.	is discussed in Section 5.2.6 iii) of the ERG report. The ERG refers the reader to the summary of errors and violations in the company's model in the ERG report Table 5.20), which provides a view on model reliability.

Issue 8 The ERG's unclear empirical reasoning to discredit the company's model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On pages 21, 143 and 148 of the report, it states: *"In two different model runs of the ERG base-case, the ICER varied by as much as £7,000 per QALY gained."* As a key justification for the ERG's claim that the model is "*of very limited value for the current decision problem and that results are to be interpreted with extreme caution*"	The ERG must amend this text to provide more details of: - What the ERG changed across the model runs - Changes to the pairwise ICERs versus orlistat and standard management, separately. - Details of the incremental costs and QALYs separately, as the small estimated QALY benefit for orlistat means that the ICER versus orlistat is highly sensitive to changes estimated health outcomes.	The ERG must provide the reader with more context for such statements, when they are being used as a basis for dismantling the company's model's credibility. The report states that: *"According to the ERG, the model should ideally be evaluated using at least 1,500 sampled patients."* Clearly this is misaligned with the ERG's claims that the model is *"of very limited value"* or that it is *"unfit for purpose"*. If 1,500 patients are indeed sufficient within the model, a further 500 patients in addition to the base-case 1,000 patients required to run in the model is clearly not immensely burdensome in regards to run time, and would inherently reduce the variation noted in model results.	This is not a factual inaccuracy. No amendments to the ERG report are needed. The sentence preceding the one quoted by the company explains the differences between the two model runs: "The large variation around the ICERs when different random numbers and sampled patient profiles are used is of particular concern. In two different model runs of the ERG base-case, the ICER varied by as much as £7,000 per QALY gained." See also the results Table in Chapter 6 for more details. Additionally, it should be noted that (as stated in section 5.3.1 of the ERG report) "the maximum number of patient profiles was restricted to 1,000" in the economic model submitted by the company. The ERG stated that "the model should be evaluated using at least 1,500 sampled

			patients". This estimate of the minimum number of patients was based on the company's diagnostic exercise, although the ERG remains uncertain that a number of 1,500 sampled patients would be nearly sufficient. These concerns are further discussed in the ERG report Section 5.2.10 in comparison to the model by Ara et al. (used by the company as a starting point), which "used a cohort of 1,000,000 patients in their patient-level simulation and stated that, with a cohort size of 200,000 patients, there was still a small amount of variation in results, which stabilised after simulation of 400,000 patients." The ERG refers the reader to this section for further detail.

Issue 9 Wording around critique of company's assumptions for weight regain

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 17 of the report, it states: *"It should be noted that the company deviated from the*	We propose the ERG amend their statement to the following: *"It should be noted that the company deviated from the base-case model assumption*	The report by Ara *et al.* states: *"It is possible that the regain would be larger than the absolute reduction achieved by the intervention, and that individuals would regain*	This is not a factual inaccuracy. No amendments to the ERG report are needed. No arguments for which the company's

assumption made by Ara et al., that patients would have regained weight to obtain the baseline BMI within three years in a linear fashion and assumed instead that patients would have regained weight to obtain the age/sex predicted BMI in three years. The company did not provide justification for why their deviation from Ara et al.'s assumption was 'logical' and plausible. Hence, to be consistent with Ara et al., the ERG preferred to assume weight regain to the baseline BMI in its base-case." This statement is inaccurate and misleading. We have clearly explained both the source and justification of our assumption to the ERG.	*made by Ara et al., that patients would have regained weight to obtain the baseline BMI within three years in a linear fashion and assumed instead that patients would have regained weight to obtain the age/sex predicted BMI in three years – a scenario analysis undertaken by Ara et al. The company provided some justification for their deviation from Ara et al.'s base-case assumption. However, to be consistent with the base-case assumption used by Ara et al., the ERG preferred to assume weight regain to the baseline BMI in its base-case."*	*more weight than they lost."* In order to incorporate weight regain appropriately, we considered it more realistic to incorporate weight regain in this manner as part of our base-case assumptions. We stated in Table 53 of the CS that: *"BMI was assumed to revert to the natural history model predicted BMI given the intrinsic correlation known between age and BMI (as shown by the natural history model in Section 5.3.4.3)."* Given that we typically observe an increase in BMI as patients age, we considered this setting as more appropriate than assuming patients refer back to their baseline BMI (derived at least 3 years prior). We also provided justification for this assumption in response to clarification question B1a. which in summary stated that simulated patient weight upon model entry was consistent with the natural history model, and therefore to ensure consistency with later BMI measures, we opted to use the same natural history model to predict BMI following weight regain. We acknowledge that the ERG may prefer their base-case setting, but consider it important to acknowledge the full reasoning why regain was assumed to follow the trajectory of the natural history model over a 3-year period. No impact is noted on model results, but clarification should aid committee understanding of why the assumption was applied.	approach should be preferred over Ara et al.'s base-case assumption have been provided by the company. It remains unclear to the ERG "why their deviation from Ara et al.'s assumption was 'logical' and plausible", as was stated in the ERG report Section 1.5, citing the company's response to the clarification letter.

Issue 10 The ERG's lack of consideration regarding the implementation of further treatment

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 18 of the report, it states: *"One major limitation of the model is the inability to incorporate re-treatment, behaviour modification treatment and/or bariatric surgery (for which patients become eligible over time once their BMI is/increases to >40kg/m^2 in the model)."* Considering the lack of intensive behavioural modification and bariatric surgery at later times as a *"major limitation"* of the model is misleading.	We propose the ERG amend their statement to the following: *"The model does not incorporate re-treatment, behaviour modification treatment and/or bariatric surgery (for which patients become eligible over time once their BMI is/increases to >40kg/m^2 in the model). The company justified the omission of further treatment in the model based on data availability and additional model complexity. However, the company noted in response to clarification question B3b that due to the differential mechanisms of action between orlistat and NB32, retreatment could be plausible in clinical practice."*	There are no data to inform model assumptions regarding subsequent treatment effectiveness. One of the key strengths of the *de novo* model is its use of GPRD data to inform TTE predictions – to incorporate further lines of treatment would not be possible without introducing additional modelling assumptions, and would not be based on comparable, extensive statistical analysis. Incorporating multiple lines of treatment would inherently be highly complex, and would therefore introduce further issues with run-time – taking bariatric surgery as an example would involve at least two prior lines. As model run time has been considered as a limiting factor of the *de novo* model, it is important to highlight that suggesting to increase model complexity (and run time) by introducing further lines of treatment based on no evidence is inappropriate and misleading.	This is not a factual inaccuracy. No amendments to the ERG report are needed.

Issue 11 Inclusion of parameters in sensitivity analysis

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 19 of the report, it states:	We propose the ERG amend their	The natural history model parameters	This is not a factual inaccuracy.

"…the ERG believes the PSA results in the CS are flawed for multiple reasons… …3) the exclusion of key input parameters from the PSA (e.g. TTD, natural history of BMI model, obesity-related events)." This is incorrect – the natural history of BMI model is reflective of patient variability (given that it is used to derive BMI at baseline and post weight regain). Therefore, the natural history model was included in all models runs (including the PSA). The time to obesity-related events parameters were omitted from the PSA as appropriate means of sampling from the time to event models were unobtainable – this statement is misleading, as it suggests that appropriate data were excluded. The 95% confidence intervals reported by Ara *et al.* do not consider the correlation between parameters, and therefore if utilised in PSA or incorrectly represent uncertainty around model results.	statement to the following (introduced wording underlined): *"…the ERG believes the PSA results in the CS are <u>limited</u> for multiple reasons… …3) the <u>unavailability of uncertainty distribution data for some</u> input parameters (e.g. TTD, obesity-related events)."*	were varied in all model runs. This is evident in the *de novo* model on the "DICE equations" sheet in cells U117:Z124. In these cells, random numbers are used to elicit draws of the coefficients used in the natural history model for each patient run (regardless of whether this in PSA or standard deterministic analysis). We acknowledge that the time to obesity-related events parameters were omitted from the PSA; however, these parameters were requested from the authors of the study our model is based on, and true to (Ara et al). We consider this is important to acknowledge in the ERG report, as the current wording suggests omission of these parameters was intentional.	No amendments to the ERG report are needed. The ERG was aware that variability of the BMI model (i.e. first order uncertainty) was reflected in the company's deterministic analysis. However, the ERG listed parameters for which second order uncertainty was not reflected in the PSA. This second order uncertainty (i.e. parameter uncertainty) of the estimated BMI is not incorporated in the PSA. Uncertainty about time to obesity-related events should also have been reflected in the PSA, despite the lack of data on correlation between parameters (e.g. without incorporating correlations between these parameters).
On page 132 of the report, it states: *"For reasons described in the previous paragraph, the ERG believes the PSA results are flawed. As a result the CEACs should be interpreted with extreme caution. The company did not perform deterministic sensitivity analyses on all parameters that are uncertain. Most notably, some*	We propose that the ERG's reporting of the deterministic sensitivity limitations would be more fairly and helpfully limited to the following: *"The correlated parameters for which the appropriate uncertainty distribution data were not*	The ERG's reporting is inflammatory and misleading in presenting one limitation (the unavailability of appropriate uncertainty data for some input parameters) as several company-driven problems (leaving parameters out of both probabilistic and deterministic sensitivity analyses),	This is not a factual inaccuracy. No amendments to the ERG report are needed.

parameters that were left out of the PSA were also not varied in deterministic sensitivity analyses, such as TTD, the probability of obesity related events, and the BMI natural history model. For instance the uncertainty around TTD, influencing both treatment effects and costs, is likely to significantly affect model results. The subgroup analyses with T2DM and non-T2DM patients should be interpreted with caution, because in these subgroup analyses the baseline characteristics (which impact obesity-related comorbidities and utility values) are independent on T2DM status. As stated in section 5.2.3 this leads to counter-intuitive patient profiles."	*available were not included in probabilistic or deterministic sensitivity analyses. That uncertainty around these parameters is not captured means that parameter precision uncertainty around model results is underestimated."*	contributing to "flawed" results. Parameter uncertainty data were not available for some parameters. This meant we could appropriately account for parameter precision uncertainty using either deterministic or probabilistic uncertainty analyses. The inclusion of model parameters that are inherently correlated in deterministic one-way sensitivity analysis can itself be considered flawed, as the correlation between parameters cannot be explored.	

Issue 12 The ERG's description of the company not providing privately owned model validation documentation as "troublesome"

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 19 of the report, it states: *"In this light, the ERG considers it troublesome that the company did not provide the results of the internal validation it performed (as requested in response to*	We propose the ERG amend their statement to the following: *"The company could not provide the results of the internal validation it performed (as requested in response to clarification question B19), as the checklist itself is commercial property. The ERG note that*	We provided details of the internal validation checks in response to clarification question B19, which stated that the checklist used is commercial property and therefore we were unable to provide the completed checklist. Considering this as *"troublesome"* is misleading, as the checklist legally could not be provided – this is not the same as suggesting the checklist was intentionally not	This is not a factual inaccuracy. No amendments to the ERG report are needed. Commercial property could be marked as commercially in confidence, hence the ERG considered the argument for not providing the results of the internal validation as weak.

clarification question B19).” This statement is misleading, as the reason for not providing the checklist is not presented.	*provision of this completed checklist may have aided internal validation.”*	provided.	

Issue 13 The ERG’s interpretation of the conclusions from systematic review of previous economic evidence

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
On page 88 of the report, it states: *“Since the identified studies did not consider NB32 as an intervention, the ERG agrees that no specific conclusion could be drawn from this review.”* This statement is factually inaccurate. The review highlighted the key relevant study considered throughout this appraisal - Ara *et al.*	We propose the ERG amend their statement to the following: *“Since the identified studies did not consider NB32 as an intervention, the ERG agrees that no specific conclusion regarding the cost-effectiveness of NB32 could be drawn from this review.”*	The study by Ara *et al.* is clearly identified within the review of previous economic analyses as a high-quality study that utilised systematic approach to search, appraise and synthesise evidence, with the stated aim of evaluating the clinical and cost-effectiveness of pharmacological treatments for overweight or obese patients, from the perspective of the UK NHS and Personal Social Services. It is important to acknowledge that though no previous NB32 cost-effectiveness studies were identified, the study by Ara *et al.* was obtained and is highly relevant to this appraisal.	This is not a factual inaccuracy. No amendments to the ERG report are needed. The quote by the company in the far left column should be considered in the context of the preceding sentence. The preceding sentence indeed indicates that “no specific conclusion” relates to the cost effectiveness of NB32. See full quote below (ERG report section 5.1.4): “The CS provides an overview of the included studies but no specific conclusion regarding the cost effectiveness of NB32, or other pharmacological treatments, is formulated. ERG comment: Since the identified studies did not consider NB32 as an

			intervention, the ERG agrees that no specific conclusion could be drawn from this review."

Issue 14 Generalisability of the trials to clinical practice

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 12 of the ERG report, and throughout: *'there are a number of limitations [of the four main trials] when applying them to clinical practice'* These are detailed as: • Little data on ethnic groups relevant to the UK (particularly people from Asia) • Few overweight as opposed to obese participants • Majority of participants are female • No patients in the main trials have previously taken orlistat	We kindly request that the ERG amend this section to acknowledge the clinical validation which has been conducted and which confirms trials were a fair reflection of the average patient seen in UK NHS Practice. Clinician feedback also confirmed that having previously received orlistat is not expected to have any effect on the efficacy of NB32	Potentially misleading interpretation of the generalisability of the trials.	Not a factual error. The ERG acknowledges that the company consulted a clinical expert who concluded that "the patient population included in the trials was a fair reflection of the average patient seen in UK NHS practice." However it remains the case that certain groups are not well represented e.g. overweight patients with comorbidities and people of Asian ethnicities.
Page 38 of the ERG report: *'All the main trials compared NB32*	We request that the ERG amend this statement to make clear that in all trials, both	As all trials included standard management in both arms, the trials compared NB32 plus standard	Not a factual error.

to placebo' Also, Page 38: *'The ERG draws to the attention of the committee that no trials directly compare NB32 to orlistat as specified in the NICE scope'*	arms included a standard management arm. Furthermore, we request that the ERG highlight the CHMP's opinion that a lack of direct comparison to orlistat was acceptable due to a number of reasons.	management to standard management alone. Therefore, the four main trials all provide a comparison of NB32 to a named comparator of interest. In addition, active comparator trials to orlistat were not considered appropriate as the distinct tolerability profile or orlistat makes it difficult to blind. Using orlistat as an active reference could have led to un-blinding of patient treatment allocation and potentially to disparate patient withdrawal patterns. The omission of an active reference was deemed acceptable by the CHMP.	

Issue 15 Minor factual Inaccuracies

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 11 of the ERG report and throughout: *'approximately 2% of participants were overweight and 98% obese'*	Across NB32 and control arms, a mean of 3% of patients were overweight and 97% were obese (note: data for the COR-BMOD study is taken from the all randomised population; data for other studies is from the full analysis set).	Minor factual inaccuracy	No change required. The ERG calculated the percentages overweight (BMI < 30 kg/m^2) using the Clinical Study Reports provided and table 15 of the submission. The actual percentage overweight was calculated to be 2.53%. In the report we use the words 'approximately 2%' Using 3% would not change our

<table>
<tr><td></td><td></td><td></td><td>overall conclusions on the low numbers of overweight patients across the trials.</td></tr>
<tr><td>Page 13 of the ERG report:

‘In the NB32 trials, 21.9% of patients receiving NB32 were randomised but excluded from the analyses’</td><td>Across NB32 arms, 18.9% of randomised patients were excluded from mITT analyses, as summarised below:
<table><tr><th></th><th>COR-I</th><th>COR-II</th><th>COR-BMOD</th><th>COR-DM</th></tr><tr><td>ITT</td><td>583</td><td>1001</td><td>591</td><td>335</td></tr><tr><td>mITT</td><td>471</td><td>825</td><td>482</td><td>265</td></tr><tr><td>% Δ</td><td>18.5</td><td>17.6</td><td>18.4</td><td>20.9</td></tr></table></td><td>Minor factual inaccuracy</td><td>Not a factual error. The discrepancy is in the number of patients analysed in the COR-II trial. As shown in Tables 31 and 32 (CS, pages 124-126), only 702 NB32 patients were included in the ITC analyses based on mITT data (rather than the 825 mentioned in the Table).</td></tr>
<tr><td>Page 23 of the ERG report:

‘The statement that women are more likely to be obese is incorrect’</td><td>A reference for a 2016 House of Commons briefing paper was provided for this statement and states in a summary box on Page 3:

‘Men are more likely to be overweight but women are more likely to be obese’</td><td>Minor factual inaccuracy</td><td>The ERG referred to a different source. However the source of the company’s statement was correct. The ERG has removed the sentence “The statement that women are more likely to be obese is incorrect. Twenty-seven percent of both genders are obese.”</td></tr>
<tr><td>On page 129 of the report, it states:

“The ERG asked for more information on disaggregated outcomes of the model, such as costs associated with events, and time with events, but these were not provided by the</td><td>We propose the ERG amend their statement to the following:

“The ERG asked for more information on disaggregated outcomes of the model, such as costs associated with events, and time with events, but these <u>could</u> not <u>be</u> provided by the company.”</td><td>A clear limitation of the chosen model structure is that the outputs are not equivalent to those that would be obtained through a standard area under the curve model. Additional non-essential outputs (such as specific cost items) would require additional</td><td>This is not a factual inaccuracy. No amendments to the ERG report are needed.

For DES models, disaggregated model outcomes can be provided and this has been done</td></tr>
</table>

company." Disaggregated outcomes are not presented as a direct consequence of the chosen model structure. The statement from the ERG is misleading, as it does not acknowledge why disaggregated results are not presented.		coding and therefore further computational burden. We would expect the ERG to be flexible enough to acknowledge and deal with the implications of alternative model structures where appropriate – failure to do so is potentially misleading.	previously (e.g. NICE TA387). In this specific case the outcomes could, for instance, be disaggregated based on events experienced (e.g. before / after treatment discontinuation, weight regain, experiencing T2DM and/or experiencing cardiovascular events).
Typographical errors in tables 4.8; 4.9; 4.10; 4.11 and 4.15	Table 4.8: SE for percentage change from baseline at end of study for NB32 in COR-DM should be 0.34 instead of 0.7 Table 4.9: SE in mITT for percentage change from baseline at end of study for NB32 in COR-DM should be 0.34 instead of 0.7 Table 4.10: in ITT imp analysis, in COR-BMOD % patients with >5% decrease in weight vs placebo should be 2.37 (1.48, 3.80) rather than NR Table 4.11: COR-1 percentage change from baseline at end of study, Pbo should read -1.85(0.66) rather than -1.83(5.92) Table 4.15: COR-DM Pbo for constipation should have SE of 7.1 rather than 17.1 Table 4.15: COR-DM NB32 for dizziness should be 39 rather than 390 Table 4.15: COR-II NB32 for vascular disorders has	Minor factual inaccuracies	This has been changed to 0.3 as in Table 24, CS, page 102. As above This figure was not provided in the CS but appears to refer to data in table 15 of the appendix describing COR-DM not COR-BMOD. Therefore it has not been added. This has now been corrected. This has now been corrected. This has now been corrected.

	missing value of 69 Table 4.15: COR-II Pbo for vascular disorders has missing value of 16 Table 4.15: COR-II NB32 for hypertension has missing value of 19 Table 4.15: COR-II Pbo for hypertension has missing value of 8		This has now been added. This has now been added. This has now been added. This has now been added.

in collaboration with:

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUTE OF HEALTH POLICY & MANAGEMENT

Maastricht University

Naltrexone-bupropion for managing overweight and obesity

ERRATUM

This document contains errata in respect of the ERG report in response to the company's factual accuracy check.

The table below lists the page to be replaced in the original document and the nature of the change:

Page nr:	**Change:**
23	Sentence removed
46	Table 4.8: SE for percentage change from baseline at end of study for NB32 in COR-DM should be 0.34 instead of 0.7
47	Table 4.9: SE in mITT for percentage change from baseline at end of study for NB32 in COR-DM should be 0.34 instead of 0.7
50	Table 4.11: COR-1 percentage change from baseline at end of study, Pbo should read -1.85(0.66) rather than -1.83(5.92)
55-56	Table 4.15: COR-DM Pbo for constipation should have SE of 7.1 rather than 17.1 Table 4.15: COR-DM NB32 for dizziness should be 39 rather than 390 Table 4.15: COR-II NB32 for vascular disorders has missing value of 69 Table 4.15: COR-II Pbo for vascular disorders has missing value of 16 Table 4.15: COR-II NB32 for hypertension has missing value of 19 Table 4.15: COR-II Pbo for hypertension has missing value of 8

ERG comment: The ERG checked the references provided to support the statements in the submission. In general these were found to be appropriate. However the ERG noted a number of discrepancies:

- Although BMI measures of overweight and obesity cited in the CS match NICE guidelines,[5] the guidelines also emphasise that BMI should be interpreted with caution and that waist circumference in people with a BMI < 35kg/m2 should be considered. The guidelines also state that "*The use of lower BMI thresholds (23 kg/m2 to indicate increased risk and 27.5 kg/m2 to indicate high risk) to trigger action to reduce the risk of conditions such as type 2 diabetes, has been recommended for black African, African-Caribbean and Asian (South Asian and Chinese) groups*."[5]
- It was unclear how exactly numbers of adults who are overweight or obese with weight-related comorbidities in England quoted in the CS were derived. No source was cited for the estimated 16% with a weight-related comorbidity.
- Women are more likely to be morbidly obese (BMI>40) than men (3.6% vs 2.2%) 68% of men were overweight or obese in 2015 compared to 58% of women.[3]
- Important variations for the prevalence of obesity have also been linked with social class. It has been suggested that this is associated with the degree of relative social inequality.[4]
- The studies supporting the link between excess weight and depression report an association only for those who are severely obese and/or have a chronic health condition.
- The report cited by the company on deaths associated with obesity referenced data collected in 2001.[3] According to the World Health Organisation, an estimated 9.6% of deaths among men and 11.5% of women are due to overweight and obesity in developed countries.[6] Applying these to England (2001 data) gives 52,500 not 34,100 deaths attributable to obesity as cited by the company.

2.2 Critique of company's overview of current service provision

The CS notes that in England *"Treatment is based upon a patient's BMI and what, if any, comorbidities are present, as outlined in Table 8"* (duplicated below).[1]

4.2.4 Direct evidence: Efficacy results

The main results of the modified intention-to-treat analysis presented in the CS are shown in Table 4.8. Tables 4.9 and 4.10 compare the mITT results for the primary outcomes with the two methods of ITT analysis **(**weight regain imputation method and using baseline-carried forward analysis).

Table 4.8: Main results of NB32 trials (mITT analysis)

	COR-I[25]		**COR-II[27]***		**COR-BMOD[26]**		**COR-DM[28]**	
	NB32	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
N	471	511	825	456	482	193	265	159
Baseline weight, mean kg (SD)	100.2 (16.3)	99.3 (14.3)	100.7 (16.7)	99.3 (16.0)	100.7 (15.4)	101.9 (15.0)	104.2 (18.9)	105.0 (17.1)
End of study weight, mean kg (SD)	94.2 (17.4)	98.0 (15.2)	94.2 (17.6)	97.2 (16.2)	91 (17.1)	96.4 (17.1)	101.0 (19.7)	103.0 (17.3)
Percent change from baseline at end of study, LS mean (SE)	-6.1 (0.3)	-1.3 (0.3)	-6.5 (0.2)	-1.9 (0.3)	-9.3 (0.4)	-5.1 (0.6)	-5.0 (0.3)	-1.8 (0.4)
NB32 – placebo, Difference of LS mean	-4.8 (-5.6 to -4.0)		-4.6 (-5.2 to -3.9)		-4.2 (-5.6 to -2.9)		-3.3 (-4.3 to -2.2)	
No of patients with ≥ 5% decrease in weight, n (%)	226 (48.0)	84 (16.4)	459 (55.6)	80 (17.5)	320 (66.4)	82 (42.5)	118 (44.5)	30 (18.9)
Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	4.9 (3.6 to 6.6)		6.6 (5.0 to 8.8)		2.9 (2.0 to 4.1)		3.4 (2.2 to 5.5)	
No of patients with ≥ 10% decrease in weight, n (%)	116 (24.6)	38 (7.4)	225 (27.3)	32 (7.0)	200 (41.5)	39 (20.2)	49 (18.5)	9 (5.7)
Patients with ≥ 10% decrease in weight, NB32 vs placebo, OR (95% CI)	4.2 (2.8 to 6.2)		5.4 (3.6 to 8.0)		2.9 (2.0 to 4.4)		3.8 (1.8 to 7.9)	

Source: Section 4.7 of the CS[1]
Footnote: *Week 28 results
BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus

Table 4.9: Comparison of mITT and ITT results: percentage weight loss

Type of analysis	Outcome	COR-I[25]		COR-II[27]*		COR-BMOD[26]		COR-DM[28]	
		NB32	Pbo	NB32	Pbo	NB32	Pbo	NB32	Pbo
mITT	Percent change from baseline at end of study, LS mean (SE)	-6.1 (0.3)	-1.3 (0.3)	-6.5 (0.2)	-1.9 (0.3)	-9.3 (0.4)	-5.1 (0.6)	-5.0 (0.3)	-1.8 (0.4)
	NB32 – placebo, Difference of LS mean	-4.8 (-5.6 to -4.0)		-4.6 (-5.2 to -3.9)		-4.2 (-5.6 to -2.9)		-3.3 (-4.3 to -2.2)	
ITT using weight regain imputation method (ITT Imp)	Percent change from baseline at end of study, LS mean (SE)	-4.6 (0.3)	-1.2 (0.3)	-5.2 (0.2)	-1.9 (0.3)	NR	NR	-3.5 (0.3)	-1.7 (0.4)
	NB32 – placebo, Difference of LS mean	-3.4 (-4.1 to -2.7)		-3.4 (-3.9 to -2.8)		NR		-1.9 (-2.8 to -0.9)	
ITT using baseline-carried forward analysis (ITT BOCF)	Percent change from baseline at end of study, LS mean (SE)	-4.0 (0.3)	-0.9 (0.3)	-4.8 (0.2)	-1.5 (0.3)	-5.9 (0.4)	- 4.0 (0.6)	-3.1 (0.3)	-1.3 (0.4)
	NB32 – placebo, Difference of LS mean	-3.1 (-3.8 to -2.4)		-3.3 (-3.9 to -2.7)		-1.9 (-3.2 to -0.6)		-1.7 (-2.7 to -0.8)	

Source: Section 4.7 of the CS[1] and CS appendices[38]

Footnote *Week 28 results

BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus, Imp = weight regain imputation, BOCF = baseline observation carried forward

Table 4.11: Results in male subgroups (mITT data)

	COR-I[25, 34]		COR-II[27, 36]		COR-BMOD[26, 35]		COR-DM[28, 37]	
	NB32	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**	**NB32**	**Pbo**
N	76	80	NR		56	17	121	75
Baseline weight, mean kg (SD)	115.36 (18.58)	112.16 (14.35)			118.11 (16.09)	122.12 (18.40)	116.83 (17.72)	111.89 (16.33)
End of study weight, mean kg (SD)	109.16 (18.10)	110.09 (15.33)			107.96 (18.99)	116.94 (20.35)	111.27 (18.58)	110.09 (15.95)
Percent change from baseline at end of study, LS mean (SE)	-5.20 (0.68)	-1.85 (0.66)			-8.75 (0.93)	-4.75 (1.70)	-4.79 (0.47)	-1.51 (0.60)
NB32 – placebo, Difference of LS mean (SE)	-3.34 (0.94)				-4.00 (1.94)		-3.28 (0.77)	
No of patients with ≥ 5% decrease in weight, n (%)	29 (38.16)	16 (20.00)			39 (69.64)	7 (41.18)	51 (42.15)	10 (13.33)
Patients with ≥ 5% decrease in weight, NB32 vs placebo, OR (95% CI)	2.36 (1.14, 4.86)				3.12 (0.99, 9.80)		4.69 (2.19, 10.05)	

Source: Trial CSRs

BMOD = intensive behaviour modification; COR = Contrave obesity research; DM = diabetes mellitus

Adverse events occurred in 83.1% to 93.7% of treatment groups and 68.5% to 88.0% of placebo groups. Approximately 58% to 76% of these were attributed to the drug in NB32 groups across the trials. Serious adverse events occurred at similar rates in treatment and placebo groups across the trials. However a larger number of patients discontinued due to adverse events across the trials (19.5% to 29.4% for treatment groups vs. 9.8% to 15.4% in placebo groups).

The main specific adverse events are listed in Table 4.15.

Table 4.15: Specific adverse events (≥ 5% in at least one treatment arm of an included trial)

	COR-I[25]		**COR-II[27]**		**COR-BMOD[26]**		**COR-DM[28]**	
	NB32 (n = 573)	**Pbo (n = 569)**	**NB32/ 48 (n = 992)**	**Pbo (n = 492)**	**NB32 (n = 584)**	**Pbo (n =200)**	**NB32 (n = 333)**	**Pbo (n = 169)**
Adverse Event, n (%)								
Gastrointestinal disorders	**292 (51.0)**	**136 (23.9)**	**532 (53.6)**	**131 (26.6)**	**380 (65.1)**	**78 (39.0)**	**215 (64.6)**	**53 (31.4)**
Nausea	171 (29.8)	30 (5.3)	290 (29.2)	34 (6.9)	199 (34.1)	21 (10.5)	141 (42.3)	12 (7.1)
Vomiting	56 (9.8)	14 (2.5)	84 (8.5)	10 (2.0)	64 (11.0)	13 (6.5)	61 (18.3)	6 (3.6)
Constipation	90 (15.7)	32 (5.6)	189 (19.1)	35 (7.1)	141 (24.1)	28 (14.0)	59 (17.7)	12 (7.1)
Dry mouth	43 (7.5)	11 (1.9)	90 (9.1)	13 (2.6)	47 (8.0)	6 (3.0)	21 (6.3)	5 (3.0)
Diarrhoea	26 (4.5)	28 (4.9)	55 (5.5)	18 (3.7)	43 (7.4)	15 (7.5)	52 (15.6)	16 (9.5)
Abdominal pain upper	NR	NR	NR	NR	32 (5.5)	3 (1.5)	17 (5.1)	3 (1.8)
Infections and infestations	**203 (35.4)**	**200 (35.1)**	**359 (36.2)**	**205 (41.7)**	**188 (32.2)**	**63 (31.5)**	**121 (36.3)**	**77 (45.6)**
Upper respiratory tract infection	57 (9.9)	64 (11.2)	86 (8.7)	55 (11.2)	38 (6.5)	18 (9.0)	26 (7.8)	16 (9.5)
Sinusitis	30 (5.2)	34 (6.0)	51 (5.1)	35 (7.1)	16 (2.7)	6 (3.0)	16 (4.8)	14 (8.3)
Nasopharyngitis	29 (5.1)	31 (5.4)	82 (8.3)	40 (8.1)	36 (6.2)	15 (7.5)	28 (8.4)	23 (13.6)
Musculoskeletal and connective tissue disorders	**72 (12.6)**	**90 (15.8)**	**159 (16.0)**	**96 (19.5)**	**104 (17.8)**	**46 (23.0)**	**58 (17.4)**	**40 (23.7)**
Nervous system disorders	**167 (29.1)**	**95 (16.9)**	**326 (32.9)**	**81 (16.5)**	**263 (45.0)**	**60 (30.0)**	**129 (38.7)**	**32 (18.9)**

	COR-I[25]		COR-II[27]		COR-BMOD[26]		COR-DM[28]	
Headache	79 (13.8)	53 (9.3)	174 (17.5)	43 (8.7)	139 (23.8)	35 (17.5)	46 (13.8)	15 (8.9)
Dizziness	54 (9.4)	15 (2.6)	68 (6.9)	18 (3.7)	85 (14.6)	9 (4.5)	39 (11.7)	9 (5.3)
Tremor	12 (2.1)	1 (0.2)	35 (3.5)	3 (0.6)	34 (5.8)	2 (1.0)	22 (6.5)	4 (2.4)
Psychiatric disorders	**85 (14.8)**	**62 (10.9)**	**205 (20.7)**	**75 (15.2)**	**145 (24.8)**	**45 (22.5)**	**75 (22.5)**	**20 (11.8)**
Insomnia	43 (7.5)	29 (5.1)	97 (9.8)	33 (6.7)	51 (8.7)	12 (6.0)	37 (11.1)	9 (5.3)
Anxiety	9 (1.6)	12 (2.1)	48 (4.8)	21 (4.3)	30 (5.1)	7 (3.5)	18 (5.4)	2 (1.2)
Vascular disorders	**51 (8.9)**	**22 (3.9)**	**69 (7)**	**16 (3.3)**	**46 (7.9)**	**7 (3.5)**	**40 (12.0)**	**12 (7.1)**
Hot flush	30 (5.2)	7 (1.2)	42 (4.2)	6 (1.2)	28 (4.8)	1 (0.5)	7 (2.1)	4 (2.4)
Tinnitus	15 (2.6)	6 (1.1)	29 (2.9)	1 (0.2)	31 (5.3)	1 (0.5)	8 (2.4)	1 (0.6)
Hypertension	17 (3.5)	14 (2.5)	19 (1.9)	8 (1.6)	14 (2.4)	4 (2.0)	33 (9.9)	7 (4.1)

Source: Tables 43, 45, 47 and 49 of the CS[1] and Trial CSRs[34-37]
Footnote: Adverse event categories in bold

The main category of adverse event occurring more frequently in treatment groups across the trials was gastrointestinal disorders. Nausea, in particular, occurred frequently and more often in treatment groups. Across the trials rates of nausea ranged from 29.2% to 42.3% in treatment groups. Rates ranged from 5.3% to 10.5% in placebo groups. Vomiting, constipation and dry mouth also occurred more frequently in treatment groups although at a lower rate than that of nausea. Nervous system disorders such as headache, dizziness and tremor occurred more frequently in treatment groups.

The incidence of events of particular concern (serious cardiovascular disorders and suicidality measured on IDS) was extremely small and any differences between groups could not be ascertained in view of the small numbers in both groups.

ERG comment:

- The ERG draws to the attention of the committee the greater proportion of gastrointestinal events, particularly nausea, in NB32 groups across the trials. Although the majority of events were not serious, more participants withdrew as a result of adverse events in treatment groups. This finding is relevant to implementation of the intervention in clinical practice.
- The ERG notes that the NB-CVOT trial (described in Section 4.2.7 of the report below) was primarily designed to investigate the cardiovascular safety of NB32 in weight management. However the study was terminated earlier than originally planned (after the 50% interim analysis), after interim data were made public in a US patent (and related Orexigen security filings) and by the EMA in the Mysimba EPAR.
- A further trial on occurrence of MACE in overweight and obese patients with cardiovascular disease receiving NB32 was requested by CHMP. Based in the US, this randomised trial will