CONFIDENTIAL UNTIL PUBLISHED

Evidence Review Group Report commissioned by the NIHR Systematic Reviews Programme on behalf of NICE

Semaglutide for managing overweight and obesity

ERRATUM

Post factual accuracy check version with corrections and updated confidentiality marking

Produced by	**Southampton Health Technology Assessments Centre**
Authors	Keith Cooper, Senior Research Fellow, Health Economics Karen Pickett, Research Fellow, Evidence Synthesis Geoff Frampton, Senior Research Fellow, Evidence Synthesis Neelam Kalita, Senior Research Fellow, Health Economics David Alexander Scott, Principal Research Fellow, Statistics Emma Maund, Research Fellow, Evidence Synthesis Jo Lord, Professorial Research Fellow, Health Economics
Correspondence to	Dr Karen Pickett Southampton Health Technology Assessments Centre (SHTAC) Wessex Institute Alpha House Enterprise Road, University of Southampton Science Park Southampton SO16 7NS www.southampton.ac.uk/shtac
Date completed	23/09/2021

Source of Funding: This report was commissioned by the National Institute for Health Research (NIHR) Systematic Reviews Programme as project number NIHR135023.

Acknowledgements

We thank Dr Katarina Kos, Consultant Physician and Service Lead, Exeter Medical Obesity Service, Royal Devon and Exeter NHS Foundation Trust, and Senior Lecturer, University of Exeter for providing expert clinical advice to the project team and for reading and commenting on a draft of this report.

We would like to thank Lois Woods, Senior Research Assistant, SHTAC, for appraising the systematic literature review search strategies and searching for ongoing and recently completed studies. We also thank Dr Jo Picot, Senior Research Fellow, SHTAC, for providing a quality assurance review of the draft report.

Declared competing interests of the authors and advisors

The authors and their advisors report none.

Rider on responsibility for report

The view expressed in this report are those of the authors and not necessarily those of the NIHR Systematic Reviews Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Cooper, K; Pickett, K; Frampton, G; Kalita, N; Scott, DA; Maund, E; Lord, J. Semaglutide for managing overweight and obesity: A Single Technology Appraisal. Southampton Health Technology Assessments Centre, 2021.

Contributions of authors

Keith Cooper critically appraised the health economic systematic review, critically appraised the economic evaluation, and drafted the report; Karen Pickett critically appraised the clinical effectiveness systematic review, drafted the report, project managed the review and is the project guarantor; Geoff Frampton critically appraised the clinical effectiveness systematic review, critically appraised the indirect treatment comparison, and drafted the report; Neelam

Kalita critically appraised the economic evaluation, and drafted the report; David Alexander Scott critically appraised the indirect treatment comparison, and drafted the report; Emma Maund critically appraised the clinical effectiveness systematic review, and drafted the report; Jo Lord critically appraised the economic evaluation, and drafted the report.

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST of FIGURES

LIST OF ABBREVIATIONS

ACS	Acute coronary syndrome
AE	Adverse event
AIC	Academic in confidence
BMI	Body mass index
BNF	British National Formulary
CI	Confidence interval
CIC	Commercial in confidence
CPRD	Clinical Practice Research Datalink
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CVD	Cardiovascular disease
DSU	Decision Support Unit
EMA	European Medicines Agency
EMC	Electronic Medicines Compendium
EPAR	European Public Assessment Report
EQ-5D-3L	European Quality of Life Working Group Health Status Measure 3 Dimensions, 3 Levels
EQ-5D-5L	European Quality of Life Working Group Health Status Measure 5 Dimensions, 5 Levels
EQ-VAS	EuroQol Visual Analogue Scale
ERG	Evidence Review Group
FAS	Full analysis set
FPG	fasting plasma glucose
GI	Gastrointestinal
HbA1c	Haemoglobin A1c
HDL	High density lipoprotein
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
HTA	Health technology assessment
ICER	Incremental cost-effectiveness ratio
IBT	Intensive behavioural therapy
IPD	Individual patient level data

ITC	Indirect treatment comparison
ITT	Intent to treat
IWQOL-Lite-CT	Impact of Weight on Quality of Life-Lite Clinical Trials Version
KOL	Key opinion leader
MI	Myocardial infarction
mITT	Modified intent to treat
NGT	Normal glucose tolerance
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NR	Not reported
OSA	Obstructive sleep apnoea
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality-adjusted life year
QoL	Quality of life
QS	Quality standard
RCT	Randomised controlled trial
RR	Relative risk/risk ratio
SAE	Serious adverse event
SBP	Systolic blood pressure
SD	Standard deviation
SE	Standard error
SF-36	Short Form-36
SLR	Systematic literature review
SmPC	Summary of product characteristics
SWMS	Specialist weight management services
T2D	Type 2 diabetes
TA	Technology appraisal
TLR	Targeted literature review
TEAE	Treatment-emergent adverse event
TSD	Technical Support Document
UK	United Kingdom
US	United States
VAS	Visual analogue scale

1 EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the evidence review group (ERG) as being potentially important for decision making. It also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail. Background information on the condition, health technology, evidence and information on the issues are in the main ERG report.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

Here, and throughout our report, we refer to semaglutide 2.4 mg/week in combination with a lifestyle intervention (including increased physical activity and a reduced-calorie diet) as 'semaglutide 2.4 mg' and placebo in combination with a lifestyle intervention (including increased physical activity and a reduced-calorie diet) as 'diet and physical activity'. We refer to liraglutide 3.0 mg in combination with a reduced-calorie diet and increased physical activity as 'liraglutide 3.0 mg'.

1.1 Overview of the ERG's key issues

Table 1 Summary of key issues

Issue number	Summary of issue	Report sections
1	Decision problem target population	2.2.3 and 2.3
2	Exclusion of orlistat as a comparator	2.2.1 and 2.3
3	Exclusion of the STEP 2 trial from the CS	3.2.1
4	Exclusion of the STEP 3 trial from the CS	3.2.1
5	The ITC results are not used in the economic model	3.4.3
6	Treatment stopping rule	4.2.6
7	Assumption that all patients with non-diabetic hyperglycaemia develop type 2 diabetes after an initial cardiovascular (CVD) event	4.2.2
8	Differences in how intercurrent events are recorded across trials may impact imputation	3.4.1
9	Results from the completed STEP 5 and STEP 8 trials are expected this year	3.2.1 and 3.2.1.3

10	Treatment duration and retreatment	2.2.2 and 4.2.2.1

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are:

- The company assumes that all patients develop type 2 diabetes after an initial CVD event, whereas the ERG does not agree with this assumption.
- The ERG assumes a different natural weight increase for the population.
- The ERG prefers to include the STEP 3 trial, which the company excluded.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Table 2 reports the base case results for semaglutide 2.4 mg versus diet and physical activity in the population with a BMI ≥ 30 and at least one co-morbidity. The incremental cost effectiveness ratio (ICER) for semaglutide vs diet and physical activity is [redacted] per QALY. Table 3 reports the base case results for semaglutide 2.4 mg versus diet and physical activity and liraglutide 3.0 mg. Semaglutide 2.4 mg is [redacted] compared to liraglutide 3.0 mg [redacted]

Table 2 Company base-case results for semaglutide 2.4 mg versus diet and physical activity (list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Diet and physical activity	[redacted]	17.924	15.269				
Semaglutide 2.4 mg	[redacted]	17.957	15.361	[redacted]	0.034	0.092	[redacted]
Key: ICER, incremental cost-effectiveness ratio; LYG, life year gained; QALY, quality-adjusted life year.							

Source: Reproduced from CS Table 52

Table 3 Subgroup results for semaglutide 2.4 mg versus liraglutide 3.0mg (list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Diet and physical activity	[redacted]	17.288	14.311				
Liraglutide 3.0 mg	[redacted]	17.331	14.401	[redacted]	0.043	0.090	[redacted]
Semaglutide 2.4 mg	[redacted]	17.349	14.444	[redacted]	0.018	0.043	[redacted]

Key: ICER, incremental cost-effectiveness ratio; LYG, life year gained; QALY, quality-adjusted life year.

Source: Reproduced from CS Table 57

The model results were most sensitive to the starting BMI.

1.3 The decision problem: summary of the ERG's key issues

Issue 1 Decision problem target population

Report section	2.2.3 and 2.3
Description of issue and why the ERG has identified it as important	The company has partly focused in their decision problem on a sub-population of the population specified in the NICE scope and draft marketing authorisation: people with a BMI ≥ 30 with at least one comorbidity (the 'target subgroup'). While we consider the focus on this subgroup is acceptable, we understand that semaglutide 2.4 mg will likely be used primarily within tier 3 services. If it is most likely to be used in this context, the NICE criteria for eligibility for bariatric surgery may more suitably define the target population (BMI ≥ 35 with at least one co-morbidity or ≥ 40 with or without comorbidities, unless new onset diabetes, in which case BMI ≥ 30, or lower for people of Asian family origin).[1] We acknowledge that NICE quality standard (QS) 127 states that adults with a BMI of ≥ 30 mg/kg^2 who have not had successful outcomes in tier 2 may be referred to tier 3, but we understand that few people with a BMI of 30 to 35 are currently treated in tier 3. An analysis of the cost-effectiveness for the bariatric surgery-eligible subgroup may be appropriate and informative.
What alternative approach has the ERG suggested?	To include a scenario analysis for this subgroup to illustrate cost-effectiveness in this population.
What is the expected effect on the cost-effectiveness estimates?	We have not been able to run a scenario analysis for this proposed subgroup, as to do this we would need to know the mean starting BMI for the starting cohort for the group from the STEP 1 trial. We have run a scenario analysis for a mean starting BMI of 42.5 (which models the cost-effectiveness for people with a BMI between 40 to 45). This resulted in more favourable ICERs for semaglutide 2.4 mg in comparison to

	physical activity and diet than when lower mean starting BMI values were used. A mean starting BMI of 42.5 may approximate that likely to be seen in our suggested subgroup. If that is the case, we expect that focusing on the subgroup is likely to result in lower ICERs for semaglutide 2.4 mg.
What additional evidence or analyses might help to resolve this key issue?	Provision of an illustrative cost-effectiveness scenario analysis for the bariatric surgery-eligible subgroup. Discussion with clinical experts about the company's positioning of semaglutide 2.4 mg in the care pathway and the clinical relevance of the company's target population, the bariatric surgery-eligible subgroup and the STEP 1 trial full analysis set population, will help resolve uncertainties about the positioning of semaglutide 2.4 mg in the care pathway and which population is most suitable for decision making.

Issue 2 Exclusion of orlistat as a comparator

Report section	2.2.1 and 2.3
Description of issue and why the ERG has identified it as important	The company have excluded orlistat as a comparator from their decision problem, as it is not widely used. We agree with the company's decision. However, as orlistat is included in the NICE scope as a comparator, this may require further consideration.
What alternative approach has the ERG suggested?	We have not suggested an alternative approach, as we agree with the company's exclusion of orlistat.
What is the expected effect on the cost-effectiveness estimates?	It is unknown what effect this might have on the cost-effectiveness estimates, as this comparator has not been included in the CS.
What additional evidence or analyses might help to resolve this key issue?	Through discussion with clinical experts about the relevance of this comparator and whether or not experts consider the company's exclusion of it from the decision problem is reasonable.

1.4 The clinical effectiveness evidence: summary of the ERG's key issues

Issue 3 Exclusion of the STEP 2 trial from the CS

Report section	3.2.1
Description of issue and why the ERG has identified it as important	The STEP 2[2] trial meets the NICE scope, but the company has not included data from it in their submission. The trial compared the efficacy of semaglutide 2.4mg to placebo, both as adjuncts to a lifestyle intervention that included a reduced-calorie diet and increased physical activity, in people with a BMI ≥ 27 kg/m^2 (overweight) or BMI ≥ 30 kg/m^2 (obese) with at least one weight-related co-morbidity who had glycated haemoglobin 7-10% (53-

	86 mmol/mol) and who had been diagnosed with type 2 diabetes. We are unclear, having only spoken to one clinical expert, whether people with type 2 diabetes might be treated with the 2.4 mg dose in practice for the purposes of weight loss and maintenance. Without inclusion of this trial, there is no data in the submission on the clinical and cost-effectiveness of semaglutide 2.4 mg for people with type 2 diabetes.
What alternative approach has the ERG suggested?	We have not suggested an alternative approach, but we believe further discussion about whether or not people with type 2 diabetes will be treated with the 2.4 mg dose of semaglutide 2.4 mg is warranted.
What is the expected effect on the cost-effectiveness estimates?	We note that the difference in percentage weight change between semaglutide 2.4 mg and placebo (diet and physical activity) was qualitatively smaller in the STEP 2 trial than the STEP 1 trial. This might indicate that ICER estimates for people who have type 2 diabetes may be higher than for those with other comorbidities.
What additional evidence or analyses might help to resolve this key issue?	Through discussion with clinical experts about whether or not the semaglutide 2.4 mg dose may be used in clinical practice for the purposes of weight loss and maintenance in people with type 2 diabetes. This will resolve whether or not the STEP 2 trial should have been included in the submission.

Issue 4 Exclusion of the STEP 3 trial from the CS

Report section	3.2.1
Description of issue and why the ERG has identified it as important	The STEP 3[3] trial meets the NICE scope and we believe data from it should have been included in the submission. The trial compared the efficacy of semaglutide 2.4mg to placebo, with both interventions administered as an adjunct to intensive behavioural therapy as part of a lifestyle intervention which included a reduced-calorie diet and increased physical activity. The trial included people with a BMI of ≥ 27 kg/m^2 (overweight) or BMI of ≥ 30 kg/m^2 (obese) with at least one weight-related co-morbidity. The company argue that IBT is not standard clinical practice in the UK. We suggest that in clinical practice, standard management is variable and so it is unlikely that a trial intervention will fully reflect clinical practice. Exclusion of this trial means that not all relevant data on the clinical effectiveness of semaglutide 2.4 mg has been included in the submission.
What alternative approach has the ERG suggested?	We suggest that the STEP 3 trial should have been included in the company's systematic literature review.
What is the expected effect on the cost-effectiveness estimates?	In the STEP 3 trial, the difference in percentage change in weight from baseline between semaglutide 2.4 mg and placebo (diet and physical activity) was qualitatively smaller than in the STEP 1 trial. As such, the trial may provide a more conservative estimate of the effectiveness of semaglutide 2.4 mg, which could potentially increase the ICERs.
What additional evidence or	Provision of scenario analyses that use both the STEP 1 trial and STEP 3 trial data to compare the cost-effectiveness of

analyses might help to resolve this key issue?	semaglutide 2.4 mg with diet and physical activity, and to compare the cost-effectiveness of semaglutide 2.4 mg with liraglutide 3.0 mg.

Issue 5 The ITC results are not used in the economic model

Report section	3.4.3
Description of issue and why the ERG has identified it as important	The unadjusted or adjusted ITC results are not used to inform the economic model. Instead, a separate calculation was performed. The mean changes from baseline from the STEP 1 trial product estimand are used directly in the economic model (CS Table 21), whilst for liraglutide 3.0 mg an odds ratio from SCALE 1839 was applied to the placebo and diet and physical activity arm from STEP 1 to give the adjusted estimates for liraglutide 3.0 mg (CS Table 23). This calculation is unclear to the ERG. It is also unclear why the unadjusted ITC could not have been used in the economic model, negating the need for this ad hoc calculation. The Company note that the ITC was "not able to produce adjusted estimates for efficacy in responders (further details are provided in Appendix D)" (CS section B.3.3.1.3). However, the ERG was unable to find any reference to this in Appendix D.
What alternative approach has the ERG suggested?	We suggest including the ITC results in the economic model.
What is the expected effect on the cost-effectiveness estimates?	The relative treatment effect values currently used in the economic model are more favourable for semaglutide 2.4 mg compared to placebo (diet and physical activity) or liraglutide 3.0 mg than the values from the ITC would be. For example, the mean weight change from baseline at 1 year used in the model is -18.47% for semaglutide 2.4 mg, -2.44% for placebo plus diet and physical activity (CS Table 21) and -10.42% for liraglutide 3.0 mg (CS Table 23). This gives higher differences in favour of semaglutide 2.4 mg (-16% vs placebo plus diet and physical activity, -8% vs liraglutide 3.0 mg) than the ITC (-12% vs placebo plus diet and physical activity, -6% vs liraglutide 3.0 mg) (ITC report Table 5). Utilising the ITC results in the economic model may therefore increase the ICERs.
What additional evidence or analyses might help to resolve this key issue?	We suggest the company should include the ITC results in the economic model. If this is not possible, they should provide a clear rationale as to why. The calculation currently used to generate the liraglutide 3.0 mg estimates used in the model should also be explained.

1.5 The cost-effectiveness evidence: summary of the ERG's key issues

Issue 6 Treatment stopping rule

Report section	4.2.6
Description of issue and why the ERG has identified it as important	The company has included a stopping rule for semaglutide 2.4 mg, whereby non-responders, i.e. people who have not lost at least 5% of their initial body weight after six months of taking the maintenance dose, would discontinue treatment. The ERG notes that a stopping rule was not included within the STEP 1 clinical trial. The CS states that it is unclear whether the marketing authorisation will include a stopping rule for semaglutide 2.4 mg (CS B3.2.3.1).
What alternative approach has the ERG suggested?	The ERG has not suggested an alternative approach; however, we feel that due to the relatively large impact of this issue on model results that it warrants further discussion.
What is the expected effect on the cost-effectiveness estimates?	The CS reports an analysis where there is no stopping rule and the treatment policy estimand has been used (CS Table 56). In this scenario, the ICER increases from [redacted] per QALY to [redacted] per QALY for semaglutide 2.4 mg vs diet and physical activity.
What additional evidence or analyses might help to resolve this key issue?	Through discussion with clinical experts and publication of the marketing authorisation for semaglutide 2.4 mg.

Issue 7 Assumption that all patients with non-diabetic hyperglycaemia develop type 2 diabetes after initial CVD event

Report section	4.2.2
Description of issue and why the ERG has identified it as important	Patients with non-diabetic hyperglycaemia are assumed to develop type 2 diabetes (T2D) following an initial CVD event. Clinical advice to the ERG suggests that it is not possible to assume that all patients will develop T2D after a CVD event. Whilst this assumption was previously used in TA664, we note that the NICE committee had reservations about this assumption and there was no good evidence to determine the proportion of people who develop type 2 diabetes after a CVD event.
What alternative approach has the ERG suggested?	The ERG prefers to assume that patients with non-diabetic hyperglycaemia would not develop T2D after an initial CVD event.
What is the expected effect on the cost-effectiveness estimates?	The CS reports an analysis where patients with non-diabetic hyperglycaemia do not develop T2D immediately after a CVD event (CS Table 56). In this scenario, the ICER increases from [redacted] per QALY to [redacted] per QALY for semaglutide 2.4 mg vs diet and physical activity.

What additional evidence or analyses might help to resolve this key issue?	Through discussion with clinical experts on the validity of this assumption

The following issues identified by the ERG in the cost-effectiveness evidence where we disagree with the company (summarised in Table 48). These are not considered key issues as they only have a relatively small impact on the model results:

- Mean increase of weight by 0.106kg/m^2 (0.296 kg) per year
- Maximum age of weight increase, 66 years
- Weight decreases after attaining the maximum age for weight increase
- Cost of microvascular complication £398
- Cost of sleep apnoea £274

1.6 Other key issues: summary of the ERG's view

Issue 8 Differences in how intercurrent events are recorded across trials may impact imputation

Report section	3.4.1
Description of issue and why the ERG has identified it as important	Differences in how or whether intercurrent events are recorded between trials raise questions about how they can be consistently handled in the missing data imputation used to calculate the trial product estimand. In SCALE 1839 the company noted there was "no notion of anti-obesity rescue medication" (clarification response A16) nor any distinction between treatment discontinuation and trial withdrawal. It is unclear to the ERG whether this means rescue medications were not recorded or not permitted.
What alternative approach has the ERG suggested?	The ERG prefer the treatment policy estimate since this uses less imputation but we realise this may not be appropriate for the economic model.
What is the expected effect on the cost-effectiveness estimates?	It is unclear whether this could have impacted the economic model nor any direction of effect.
What additional evidence or analyses might help to resolve this key issue?	It is unclear how this issue could be resolved.

Issue 9 Results from the completed STEP 5 and STEP 8 trials are expected this year

Report section	3.2.1 and 3.2.1.3
Description of issue and why the ERG has identified it as important	The company has not included data from the completed STEP 5 and STEP 8 trials in the CS, as they stated data from the trials were not available in time for this submission. The STEP 8 trial was a head-to-head comparison of semaglutide 2.4 mg with liraglutide 3.0 mg and also with placebo (all as adjuncts to a lifestyle intervention) in people living with obesity (BMI ≥ 30 kg/m^2) or overweight (BMI ≥ 27 kg/m^2) with at least one weight-related comorbidity. Currently, there are no other head-to-head trials available comparing semaglutide 2.4 mg and liraglutide 3.0 mg. In the CS, the company compares the clinical efficacy of the drugs in the liraglutide-eligible subgroup using an indirect treatment comparison. The STEP 5 trial compares semaglutide 2.4 mg against placebo (both as adjuncts to a lifestyle intervention) in people living with obesity (BMI ≥ 30 kg/m^2) or overweight (BMI ≥ 27 kg/m^2) with ≥ 1 weight-related comorbidity. The drugs are administered during a 104-week period. The STEP 1 trial used a 68-week treatment period, so the STEP 5 trial will provide evidence of efficacy when it is used over a longer period. Both trials are relevant to the NICE scope for this appraisal, albeit it is unclear how many people in the STEP 8 trial might be included in a 'liraglutide-eligible' subgroup. Data from these trials could potentially have a bearing on conclusions about the clinical effectiveness and cost-effectiveness of semaglutide 2.4 mg.
What alternative approach has the ERG suggested?	None. The company states the results of these studies are not currently available and the clinical study reports are expected in Q4 (STEP 8) and Q3 (STEP 5) this year.
What is the expected effect on the cost-effectiveness estimates?	The results of these studies are not available, so it is unknown what impact they may have on the cost-effectiveness estimates.
What additional evidence or analyses might help to resolve this key issue?	Provision of the results of these trials when they are available.

Issue 10 Treatment duration and retreatment

Report section	2.2.2 and 4.2.2.1
Description of issue and why the	In their economic model, the company has assumed that people are treated with semaglutide 2.4 mg for a maximum of two years and do not receive retreatment with pharmacotherapy. We agree that these assumptions are reasonable. We note,

ERG has identified it as important	however, from a discussion with our clinical expert that there are uncertainties about whether people would receive a single course of treatment and if it could be repeated. We also note that in TA664[4] the committee discussed that limiting treatment to two years was not ideal for a long-term condition such as obesity, although the committee accepted this assumption. Treatment duration and retreatment are therefore areas of uncertainty.
What alternative approach has the ERG suggested?	We have not suggested an alternative approach, but we believe further discussion about length of treatment and whether people might be retreated is warranted.
What is the expected effect on the cost-effectiveness estimates?	The ERG conducted a scenario with the ERG's preferred assumptions using a treatment duration for 3 years. In this scenario the ICER for semaglutide 2.4mg increased from [redacted] per QALY to [redacted] per QALY.
What additional evidence or analyses might help to resolve this key issue?	Through discussion with clinical experts about length of treatment and whether it is possible that some people may be retreated with semaglutide or receive treatment beyond 2 years.

1.7 Summary of ERG's preferred assumptions and resulting ICER

Based on the ERG critique of the company's model (discussed in section 4), we have identified the following aspects of the company base case with which we disagree. Our preferred assumptions are the following:

- **Patients with non-diabetic hyperglycaemia transitioning to T2D after CVD events**: We assume that these patients do not transition to T2D after CVD events.
- **Natural weight increase:** We use a natural weight increase of 0.296 kg per year.
- **Age at weight increase:** We assume weight does not increase after age 66 years.
- **Weight increase after age 66 years:** We assume that individuals lose 0.296 kg per year after age 66 years.
- **Annual cost of microvascular complications:** We use an annual cost of £398.
- **Annual cost of sleep apnoea:** We use an annual cost of £274.

Table 4 reports the ERG preferred base case results for semaglutide 2.4 mg versus diet and physical activity in the population with a BMI ≥ 30 and at least one co-morbidity. The incremental cost-effectiveness ratio (ICER) for semaglutide 2.4mg vs diet and physical activity is [redacted] per QALY. Table 5 reports the results for semaglutide 2.4mg versus diet and physical activity and liraglutide 3.0 mg for the liraglutide-eligible subgroup. Semaglutide 2.4mg is [redacted] compared to liraglutide 3.0mg [redacted]

Table 4 ERG's preferred model assumptions

Assumption	Treatments	Total costs	Total QALYs	ICER (£/QALY)
Company base-case	Diet & physical activity	[redacted]	15.269	
	Semaglutide 2.4mg	[redacted]	15.361	[redacted]
Patients with pre-diabetes do not transition to T2D after CVD events	Diet & physical activity	[redacted]	15.329	
	Semaglutide 2.4mg	[redacted]	15.419	[redacted]
+ Mean increase of weight by 0.296 kg per year	Diet & physical activity	[redacted]	15.484	
	Semaglutide 2.4mg	[redacted]	15.582	[redacted]
+ Mean decrease in weight after age 66 years: 0.296 kg per year	Diet & physical activity	[redacted]	15.540	
	Semaglutide 2.4mg	[redacted]	15.634	[redacted]
+ Age at which weight no longer increases: 66 years	Diet & physical activity	[redacted]	15.562	
	Semaglutide 2.4mg	[redacted]	15.656	[redacted]
+ Annual cost of microvascular complication, £398	Diet & physical activity	[redacted]	15.562	
	Semaglutide 2.4mg	[redacted]	15.656	[redacted]
+ Annual cost of sleep apnoea, £274	Diet & physical activity	[redacted]	15.562	
	Semaglutide 2.4mg	[redacted]	15.656	[redacted]
ERG base case	Diet & physical activity	[redacted]	15.562	
	Semaglutide 2.4mg	[redacted]	15.656	[redacted]

Table 5 ERG's preferred model assumptions- liraglutide eligible subgroup

Assumption	Treatments	Total costs	Total QALYs	Incremental ICER (£/QALY)
Company base-case	Diet & physical activity	[redacted]	14.311	
	Liraglutide 3.0mg	[redacted]	14.401	[redacted]
	Semaglutide 2.4mg	[redacted]	14.444	[redacted]
Patients with pre-diabetes do not transition to T2D after CVD events	Diet & physical activity	[redacted]	14.419	
	Liraglutide 3.0mg	[redacted]	14.505	[redacted]
	Semaglutide 2.4mg	[redacted]	14.548	[redacted]
+ Mean increase of weight by 0.296 kg per year	Diet & physical activity	[redacted]	14.562	
	Liraglutide 3.0mg	[redacted]	14.648	[redacted]
	Semaglutide 2.4mg	[redacted]	14.690	[redacted]
+ Mean decrease in weight after age 66 years: 0.296 kg per year	Diet & physical activity	[redacted]	14.642	
	Liraglutide 3.0mg	[redacted]	14.727	[redacted]
	Semaglutide 2.4mg	[redacted]	14.770	[redacted]
+ Age at which weight no longer increases: 66 years	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]
+ Annual cost of microvascular complication, £398	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]
+ Annual cost of sleep apnoea, £274	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]
ERG base case	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

This report is a critique of the company's submission (CS) to NICE from Novo Nordisk on the clinical effectiveness and cost-effectiveness of semaglutide 2.4 mg for managing overweight and obesity. It identifies the strengths and weakness of the CS. A clinical expert was consulted to advise the evidence review group (ERG) and to help inform this report.

Clarification on some aspects of the CS was requested from the company by the ERG via NICE on 5th August 2021. A response from the company via NICE was received by the ERG on 26th August 2021 and this can be seen in the NICE committee papers for this appraisal.

2.2 Background

2.2.1 Background information on overweight and obesity

The CS (section B.1.3) provides a clear and accurate overview of obesity (BMI ≥ 30 kg/m^2), including its definition, causes, prevalence, effect on health-related quality of life (HRQoL) and the morbidity and mortality associated with it. The CS outlines some of the weight-related co-morbidities people living with obesity may experience, including prediabetes, type 2 diabetes and cardiovascular disease, acknowledging that it is not an exhaustive list due to the range of complications that exist. The company do not mention eating disorders, such as binge eating (which our clinical expert states are common in this population) and the process by which mental health co-morbidities should be addressed. The ERG's clinical expert stated that screening people for mental health issues, such as depression and eating disorders, is central to their work in weight services. They see many people with eating disorders and the majority of people living with obesity who they treat have a history of depression (up to 70%). Our expert stated that whilst there may be regional variation, other services also report a high incidence of depression and anxiety. The PHQ depression screening tool was used in the pivotal semalutide trial[5] and people were only included in the trial if they had a score of < 15 on this (see CS Table 4) (scores of 15 and 20 represent moderately severe and severe depression, respectively[6]). Our expert commented that it is unclear how this should influence clinicians' prescribing in clinical practice when using semaglutide.

The CS provides information about how weight losses of 5%, 5-10% and ≥15% can positively impact co-morbidities. We understand from our clinical expert that many people achieve a weight loss of 5% in one year when under the treatment of weight management

services. For example, one evaluation of a weight management service found that 60.0% of the participants included achieved a 5% or more weight loss at 12 months.[7]

The company's description of the health condition does not include information about overweight. As discussed in section 2.3, the company have focused their decision problem on people with obesity (BMI ≥ 30 kg/m^2) who have at least one co-morbidity and have not included people living with overweight. We consider that this is acceptable (see section 2.2.3 for further discussion about this).

CS section B.1.3.4 provides information on current service provision in the NHS in England for overweight and obesity. As outlined in the CS, care is provided through four weight management tiers (tiers 1 to 4). These are shown in CS Figure 1. The CS states the tiers are a guide only and definitions can vary locally.

The CS accurately indicates that lifestyle intervention to change people's diet and physical activity is a central part of treating obesity. The CS does not provide information about the form this typically takes in practice. We understand from our clinical expert that in tier 2 services are typically managed in primary care, although provision can vary regionally. Tier 2 lifestyle interventions may take the form of, for example, referral to an exercise scheme or commercial weight management programme. People need to have taken part in tier 2 interventions before attending tier 3 services (although tier 2 services are not universally nationally available). Our expert stated that in tier 3 services, run by multidisciplinary teams, standard management of obesity involves a full assessment of an individual's mental health (including eating disorders), co-morbidities and readiness to engage with treatment. Some patients may need mental health services/treatments first as mental health issues can be a barrier to engagement with lifestyle interventions. This is one example of a service and our expert stated that there is some variation in local pathways. After assessments of suitability and readiness of engagement in lifestyle interventions, weight loss interventions are primarily delivered by dietetic services. These usually consist of group sessions with some behavioural intervention (such as motivational interviewing). They address healthy eating, having a balanced diet and eating behaviour. Emotional eating and psychological barriers such as dealing with setbacks are discussed in these sessions, which are supervised by a psychologist. People typically take part in one or two group sessions a month over six months (typically six to nine sessions over this period). Some patients are also referred for physical activity intervention.

CS section 1.3.4.1 suggests the main aim of tier 3 is to achieve clinically meaningful weight loss, and that another part of its aim is to prepare some selected patients for bariatric surgery. Expert advice to the ERG is that a key purpose of tier 3 services is to assess people's readiness for weight loss (bariatric) surgery and to prepare them for this. If people chose to undergo surgery, surgical referral takes place around six months to a year into treatment. Our clinical expert stated that around 25% to 30% of people treated in tier 3 services progress to weight loss surgery. Prior to surgical referral, prebariatric patients may require additional psychological assessment to ensure they have adequate coping mechanisms to undergo the surgical route.

The CS accurately outlines that the only pharmacological treatments currently available for people with obesity are orlistat and liraglutide 3.0 mg. The CS states NICE recommends liraglutide 3.0 mg for people with a BMI of ≥ 35 kg/m^2 who have non-diabetic hyperglycaemia and a high risk of cardiovascular disease.[4] We additionally note that liraglutide 3.0 mg is recommended by NICE for members of some minority ethnic groups at a lower BMI threshold of 32.5 kg/m^2. The recommended population in the NICE guidance is a subpopulation of the people in whom liraglutide 3.0 mg is recommended in its marketing authorisation.[8] It is indicated for people with a BMI of ≥ 27 kg/m^2 to < 30 kg/m^2 (overweight) with at least one weight-related comorbidity or people with a BMI ≥ 30 kg/m^2. The company (Novo Nordisk) markets liraglutide 3.0 mg.

The company outline that orlistat is not widely used, and that many people decide not to use it or stop taking it due to undesirable side effects. We understand from our clinical expert that orlistat has undesirable gastrointestinal side effects and that it is not used in specialised services but is still prescribed by some GPs. We also note that clinical experts informed the liraglutide 3.0 mg, TA664 appraisal committee that many people decide not to take orlistat or cease treatment with it due to side effects.[4] In the liraglutide 3.0 mg appraisal, the experts stated that most people who are referred to tier 3 services will have previously been treated with orlistat. The committee concluded that orlistat was not an alternative treatment to liraglutide 3.0 mg.[4]

2.2.2 Background information on semaglutide

The company describe semaglutide in CS section B.1.2. Semaglutide is a GLP-1 analogue that has effects on areas of the brain involved in regulation of food intake. The maintenance dose for treating overweight and obesity is 2.4 mg/week and the company refer to the intervention specifically as 'semaglutide 2.4 mg' throughout the CS (they indicate that they

do this to distinguish it from its diabetes indication). The CS states the semaglutide 2.4 mg marketing authorisation application was submitted to the EMA on 18 December 2020, with the result expected on 22 January 2022 (CS Table 2). The company provided the draft summary of product characteristics (SmPC) with the submission (CS Appendix C).

In line with the draft SmPC, the CS states that semaglutide 2.4 mg is indicated as an adjunct to a reduced-calorie diet and increased physical activity for weight management, including weight loss and maintenance, in adults with a BMI of:

- ≥30 kg/m^2 (obesity), or
- ≥27 kg/m^2 to <30 kg/m^2 (overweight) who have at least one weight-related co-morbidity.

It is self-administered once-weekly by subcutaneous injection. The dose is escalated over a 16-week period to reach a maintenance dose of 2.4 mg once weekly.

The company state in the CS that it is unclear whether the marketing authorisation will include a stopping rule for semaglutide 2.4 mg. As outlined in CS section B.1.1, the draft SmPC states that if people have not lost at least 5% of their initial body weight after six months of taking the maintenance dose, a decision should be made about whether or not to continue treatment, based on the risks and benefits to the individual person. The company applied this stopping rule to semaglutide 2.4 mg treatment in their CS base case economic model, and also conducted a scenario analysis in which no stopping rule was applied (CS section B.3.3.4.1). In both the base case and scenario analysis, the comparator's (liraglutide's) stopping rule was applied. The ERG's clinical expert commented that there is also a question about whether to continue semaglutide 2.4 mg prescriptions for people who do not engage in lifestyle intervention and discontinue their engagement in tier 3 services.

The draft SmPC does not state for how long people should be treated with semaglutide 2.4 mg. In the company's CS economic model base case, they have applied a maximum treatment duration of two years (CS section B.3.3.4). The company state treatment is typically provided in weight management services for two years and that this assumption is in line with the liraglutide, TA664 appraisal.[4] We note that in TA664, the committee discussed that limiting treatment to two years was not ideal for a long-term condition such as obesity. They noted the clinical need to reduce weight and then maintain weight loss. In the end, the committee accepted a treatment duration of two years for a single course of treatment and decided that the assumption was reasonable in the context of tier 3 weight management services.[4] Based on the committee's conclusion in TA664 and advice from our

clinical expert, we consider it is also a reasonable assumption for treatment with semaglutide 2.4 mg. Our clinical expert noted, however, that it is currently unclear what would happen regarding pharmacological treatment with liraglutide 3.0 mg or semaglutide 2.4 mg after two years. For example, it is unclear if people should receive a single course of treatment with semaglutide 2.4 mg or whether, and when, it could be repeated.

2.2.3 The position of semaglutide 2.4 mg in the treatment pathway

The company detail their proposed positioning of semaglutide 2.4 mg in the clinical care pathway in CS section B.1.3.4. The company appear to suggest (we found the text to be unclear) that semaglutide 2.4 mg would be used in tier 3 and 4 multidisciplinary team weight assessment and management clinics, where pharmacotherapy can be provided under the guidance of such a team. The company refer to these settings as specialist weight management services (SWMS). The CS states there is a need for additional pharmacological treatments within SWMS. It states orlistat is rarely used and there is an unmet clinical need for people who would not be eligible for liraglutide 3.0 mg (which is recommended by NICE for people with a BMI ≥ 35 kg/m^2 [or ≥ 32.5 kg/m^2 for members of some minority ethnic groups] who have non-diabetic hyperglycaemia and a high risk of cardiovascular disease). They state semaglutide 2.4 mg should be used as an adjunct to a reduced-calorie diet and increased physical activity in people with a BMI of ≥ 30 mg/kg^2 and at least one weight-related comorbidity. The company suggest this population is anticipated to benefit the most within SWMS from pharmacological treatment.

As noted in section 2.3, the population of people with a BMI of ≥ 30 mg/kg^2 and at least one weight-related comorbidity in whom the company proposes semaglutide 2.4 mg will be used is narrower than the draft marketing authorisation indication and the population defined in the NICE scope. The proposed population also does not fully match the population we understand to be eligible for treatment within SWMS. We note from published reports and our clinical expert that SWMS are usually provided for people with a BMI of ≥ 35 with co-morbidities or of ≥ 40 with or without comorbidities.[9 10] The company mention that NICE quality standard (QS) 127 states that adults with a BMI of ≥ 30 mg/kg^2 who have not had successful outcomes from tier 2 services should be offered a discussion about alternative weight management interventions, including referral to SWMS (i.e. tier 3). We acknowledge that NICE QS 127[11] states this, but we understand from our clinical expert that people with a BMI of 30 to 35 are currently only treated in tier 3 services if they have new onset diabetes and are preparing for weight loss surgery, which is in line with the NICE pathway for referral of people suitable for bariatric surgery into tier 3 and onwards.[1] Our clinical expert stated that

if semaglutide 2.4 mg were to be recommended for people with a BMI of ≥ 30 with other co-morbidities, this would expand the patient population for tier 3 services, which would result in additional costs.

The company have also not explicitly explained why they have not positioned semaglutide 2.4 mg as a tier 2 pharmacological intervention, as well as one that can be used in tiers 3 and 4. Clinical expert advice to the ERG is that lifestyle interventions would need to be deliverable when using semaglutide 2.4 mg and specialist assessments made. Given that the company expect semaglutide 2.4 mg to be used in SWMS, based on clinical expert advice we consider it is reasonable in this context for the company not to have positioned semaglutide 2.4 mg as an intervention for people with a BMI between 27 and 30 who have at least one weight-related co-morbidity (part of the population of interest specified in the NICE scope and draft SmPC), as they are not treated in tier 3 services.

Overall, we suggest that the company's positioning of semaglutide 2.4 mg as a treatment specifically for people with a BMI ≥ 30 who have at last one weight-related comorbidity is acceptable, if it is to be used in the NHS only within SWMS. It should be acknowledged, though, that most people who are seen in these services will have a BMI of ≥ 35 – few people currently treated within these services will have a BMI of 30 to 35.

We note that whilst the company have set out that there is an unmet need for other pharmacological treatment options within SWMS, they have not outlined in the CS how treatment with semaglutide 2.4 mg may potentially fit with weight loss surgery in the clinical pathway. It is unclear from the CS when weight loss surgery would be offered to people taking semaglutide 2.4 mg. Our clinical expert indicated that pharmacological treatment options becoming available may mean that some people may wish to try weight loss drugs before having surgery. This may make it difficult to assess people's readiness for surgery, as it will be less clear how prepared people are to change lifestyle behaviours than when treated by standard management alone. Clinical expert advice to the ERG is that, in their opinion, semaglutide 2.4 mg should be positioned for people who are eligible for but will not consider surgery or who are not fit enough to undergo it, as well as for those with a BMI of 30 to 35 with comorbidities. If people take semaglutide 2.4 mg and then decide they wish to have surgery after all, a reasonable aim could be that their weight remains stable for six months after ceasing the maintenance semaglutide 2.4 mg dose before being referred for surgery. Our expert noted that a time interval between completing pharmacological therapy with a GLP-1 analogue and commencing the surgical pathway would aid pre-bariatric surgery assessment.

ERG conclusion

The company's positioning of semaglutide 2.4 mg as a treatment specifically for people with a BMI ≥ 30 who have at last one weight-related comorbidity is acceptable, if it is intended that semaglutide 2.4 mg will only be provided in the NHS in tier 3 and 4 services (we note it is most likely to be used within tier 3). We note, though, that this would expand the patient population typically treated in tier 3 services to include more people with a BMI of 30 to 35. Currently few people with a BMI in this range are treated within tier 3 services. The company's assumption that maximum treatment duration with semaglutide 2.4 mg would be two years appears reasonable, given the precedence set by the liraglutide appraisal (TA 664),[4] but it is unclear if people would receive a single course of treatment or whether, and when, it could be repeated.

2.3 Critique of the company's definition of the decision problem

Table 6 compares the company's decision problem to the final scope for this appraisal issued by NICE. The ERG consider that the decision problem adheres to the NICE scope with the following exceptions:

- **Population**:
    - The population specified in the NICE scope and the anticipated marketing authorisation (provided by the company in CS Appendix C) is adults who have a BMI of ≥30 kg/m^2 (obese) or a BMI of ≥27 kg/m^2 to <30 kg/m^2 (overweight) in the presence of at least one weight-related comorbidity. The company have focused their submission on a narrower population: namely, adults who have a BMI of ≥30 kg/m^2 and at least one weight-related comorbidity.
    - We believe it is acceptable to focus on this subgroup, as it is still within the draft SmPC indication, the NICE scope and, as we concluded in section 2.2.3, it to some extent reflects the people who are typically treated within SWMS (where the company appears to be positioning semaglutide 2.4 mg in the clinical pathway). We understand, however, that few people with a BMI of 30 to 35 are currently treated in SWMS. Focusing on this subgroup is inclusive of these few, but overall, we consider data on the clinical efficacy of semaglutide 2.4 mg in people who have a BMI of ≥ 35 might be more representative of the clinical effectiveness likely to be achieved in practice. In this regard, the NICE criteria for eligibility for bariatric surgery may more suitably define the target

population (BMI ≥ 35 with at least one co-morbidity or ≥ 40 with or without comorbidities, unless new onset diabetes, in which case BMI ≥ 30, or lower for people of Asian family origin).[1] We understand from our clinical expert that this is the patient group that is typically treated in tier 3. Although, we acknowledge that trial data is only available for people who had co-morbidities.

  - Regarding comorbidities, efficacy evidence for people with diabetes as a comorbidity is not included (we discuss this further in section 3.2.1). Clinical expert advice to the ERG is that they expect semaglutide 2.4 mg to be used to treat overweight and obesity in people who have type 2 diabetes as a comorbidity. The company state in CS Table 14 that this population is not relevant to the submission, but they do not explain why. NICE and the ERG sought clarification from the company about the reason for this. In clarification response A1, the company explained that semaglutide 2.4 mg could potentially be used in people living with type 2 diabetes, but clinical expert advice suggested that treatment for this group would typically follow a diabetes treatment pathway where semaglutide is indicated at a lower dose. We understand from our clinical expert that, in this context, semaglutide would be used without specialist lifestyle interventions as offered in tier 3 services. The expert stated that diabetes specialists would need to refer people to obesity services for lifestyle intervention if semaglutide 2.4 mg were to be used for the management of weight at the highest dose. We suggest that, overall, it is unclear if semaglutide 2.4 mg in combination with lifestyle intervention might be used for weight loss or management in some people with type 2 diabetes. We have only been able to obtain one expert's opinion about this. It is therefore unclear if data relating to this population should have been included in the CS.

- **Comparators.** The company have excluded orlistat as a comparator, as it is not widely used. As we outlined in section 2.2.1, we understand that orlistat is not typically used in tier 3 services. We therefore consider it is reasonable for the company to have excluded it as a comparator, given the company appears to be positioning semaglutide 2.4 mg as a treatment option within SWMS.

Table 6 Summary of the decision problem

Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comments
Population			
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 at least one weight-related comorbidity	Adults who have a BMI of ≥ 30 kg/m^2 (obese) in the presence of at least one weight-related comorbidity	The company state that it is anticipated that this subgroup of people will benefit the most from pharmacological treatment within SWMS. They state that there is an unmet clinical need for this patient group, because patients have limited treatment options and many do not meet the criteria for pharmacological treatment with liraglutide 3.0 mg.	The company's focus on this subgroup is acceptable, given the company's positioning of semaglutide 2.4 mg as a treatment option within SWMS – see our discussion about this in this section and section 2.2.3.
Intervention			
Semaglutide	Semaglutide 2.4 mg	The company outline that semaglutide 2.4 mg (used as an adjunct to diet and physical activity) is an approved treatment for adults with type 2 diabetes mellitus, at doses of 0.25 mg, 0.5 mg and 1 mg. Semaglutide 2.4 mg is the specific maintenance dose for treatment of obesity.	The intervention reflects the NICE scope and is appropriate. We note the draft SmPC states that semaglutide 2.4 mg is indicated as an adjunct to a reduced-calorie diet and increased physical activity when used for weight management in people living with overweight in the presence of at least one weight-related comorbidity or living with obesity.

Comparators			
• Standard management without semaglutide (including a reduced calorie diet and increased physical activity) • Liraglutide (for the population for whom liraglutide is recommended in technology appraisal 664: patients with a BMI ≥ 35 mg/kg^2 with non-diabetic hyperglycaemia and high risk of cardiovascular disease) • Orlistat (prescription dose)	• Standard management without semaglutide (including a reduced calorie diet and increased physical activity) • Liraglutide 3.0 mg (for the population for whom liraglutide is recommended in TA664: patients with a BMI ≥ 35 mg/kg^2 with prediabetes and high cardiovascular risk)	The company state orlistat is not a relevant comparator. They suggest it is not widely used and that many people decide not to use it or stop taking it due to undesirable side effects, citing discussions held during the TA494[12] and TA664[4] appraisals and trends in prescription data.	The company's inclusion of standard management and liraglutide 3.0 mg as comparators matches the NICE scope. The company have accurately outlined the population in whom liraglutide 3.0 mg is recommended, but we additionally note that liraglutide is recommended for members of some minority ethnic groups at a lower BMI threshold of 32.5 kg/m^2.[4] We agree that orlistat is not a relevant comparator and therefore the company's exclusion of it from the decision problem is appropriate.
Outcomes			
• BMI • weight loss • waist circumference • incidence of type 2 diabetes • glycaemic status • cardiovascular events • mortality • adverse effects of treatment • health-related quality of life.	As per scope	Not applicable	Decision problem matches the NICE scope.
Economic analysis			
See CS Table 1 – text not replicated here to reduce table size	Same as NICE scope	Not applicable	The CS economic analysis has been conducted in line with the reference case stipulations outlined in the scope. The economic model base case outcomes and costs are estimated over a lifetime horizon of 40 years. Semaglutide 2.4 mg does not currently have an agreed patient access scheme (PAS)

			(discussions are ongoing with NHS England). Liraglutide 3.0 mg has a commercial access agreement and the company have provided the results of cost-effectiveness analyses with this applied.
Subgroups			
None	The submission will also address the subset of patients who are eligible to receive treatment with liraglutide 3.0 mg (patients with a BMI ≥ 35 mg/kg^2 with prediabetes and high CVD risk) following its approval in TA664.	Not applicable (specified in final scope under comparators)	Inclusion of this subgroup is appropriate.
Special considerations including issues related to equity or equality			
None	Company stated 'N/A'. We note that the company outline equality considerations in CS section B.1.4, including BMI threshold variations between different ethnicities related to their risks of developing health conditions and for intervening to prevent type 2 diabetes.	Not applicable	Clinical expert advice to the ERG is that BMI thresholds for intervention should be adjusted to take into account ethnicity, as was done in NICE's liraglutide 3.0 mg guidance.[4] Neither we nor our expert identified any other equity or equality issues.

Source: adapted version of CS Table 1

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company report three systematic literature reviews in the CS:

1. a clinical effectiveness evidence review that identified semaglutide 2.4 mg studies for inclusion in the CS and semaglutide 2.4 mg and/or liraglutide 3.0 mg studies for inclusion in an indirect treatment comparison (ITC) the company included in the CS,
2. a review of cost-effectiveness/cost-utility, costs and healthcare resource studies, and,
3. a review of HRQoL studies. A brief critique of the company's review of clinical effectiveness studies is provided in Table 7 below.

Across these reviews, we identified some concerns about the company's approach to searching for literature, study selection and the processes of data extraction and risk of bias assessment, which we detail below.

3.1.1 Searches

The CS reports three systematic searches:

- Clinical effectiveness studies (CS Appendix D.1)
- Cost-effectiveness studies, costs and resource use (CS Appendices G and I)
- HRQoL studies (CS Appendix H)

Each search had some limitations to the sources searched and search terms used (see Table 7). However, overall, the ERG consider the searches to be broadly fit for purpose, and it is unlikely that key studies have been missed. Clinical experts advising the ERG were not aware of any relevant studies that have not been identified. As the company did not search trial registries, the ERG searched ClinicalTrials.gov and the EU Clinical Trials Register for ongoing or recently completed studies of semaglutide 2.4 mg and/or liraglutide 3.0 mg. The results are discussed in section 3.2.1.3.

3.1.2 Study selection

For the cost-effectiveness and HRQoL reviews, one reviewer conducted study selection for each review, with a second reviewer checking only in cases of uncertainty (CS Appendices D.1.2, G.4 and H.3). Ideally dual reviewer screening would have been preferable to reduce the risk of errors or bias being introduced.

CS Appendix D Table 5 provides a list of studies excluded during full text screening from the clinical effectiveness review. The company did not provide the full reference citations for

these or PDFs of the references. These were requested in clarification question A2. After considering clarification response A2 we believe the reasons for excluding clinical effectiveness studies listed in CS Appendix Table 5 are appropriate.

For the clinical effectiveness review the company excluded a trial of semaglutide 2.4 mg/week which included people with type 2 diabetes who were living with overweight r obesity (STEP 2) because they considered people with type 2 diabetes would not be managed under a weight management pathway. It is unclear whether or not this trial should have been included in the CS (see further discussion in section 2.3 and section 3.2.1). The company identified another trial (STEP 3) as being eligible for inclusion in the review but excluded the trial post hoc, arguing that intensive behavioural therapy (IBT) support for diet and physical activity in the trial was not reflective of NHS practice. As explained in section 3.2.1 below, the ERG disagree with the company and believe the STEP 3 trial should have been included in the review. The ERG have no concerns with study selection in the other reviews.

3.1.3 Data extraction and risk of bias assessment

The company do not report the number of reviewers involved in the data extraction process for the HRQoL and cost-effectiveness reviews; and they do not report the number of reviewers involved in the risk of bias assessments for any of the reviews.

3.1.4 Summary of the ERG's critique

Overall, despite our concerns listed here, the company's evidence reviews are broadly fit for purpose and appear to have identified all relevant studies. However, the ERG disagree with the company's exclusion of the STEP 3 trial.

Table 7 ERG appraisal of systematic review methods

Systematic review components and processes	ERG response	ERG comments
Was the review question clearly defined using the PICOD framework or an alternative?	Yes	The PICOS is defined in CS Appendix D.1.2 table 4 for the eligibility criteria.
Were appropriate sources of literature searched?	Yes but sources could have been	The company's searches included Medline, Embase, Cochrane Central

	more comprehensive	Register of Controlled Trials, HTA Database, Database of Abstracts of Reviews of Effectiveness, but not clinical trial registries, websites, or reference lists of relevant papers or systematic reviews.
Was the time period of the searches appropriate?	Yes	Databases were searched from inception, and conferences for the past 3-4 years. Searches were updated 26th April 2021.
Were appropriate search terms used and combined correctly?	Partly	Search strategies in CS Appendix Tables 1 to 3 contain no search terms for the comparator (diet and physical activity). This is likely inconsequential as relevant RCTs would be captured by the drug search terms. However, synonyms for overweight and obesity are inadequate, and some relevant subject headings for population and comparator are missing, meaning that some relevant studies might have been missed.
(1) Were inclusion and exclusion criteria specified? (2) If so, were these criteria appropriate and relevant to the decision problem?	(1) Yes (2) Partly	(1) CS Appendix Table 4 lists the eligibility criteria. (2) BMI and HRQoL are specified outcomes in the decision problem but are not listed in the eligibility criteria. As far as the ERG are aware this did not result in the exclusion of any RCTs that would have otherwise been eligible (relevant RCTs would be captured on other PICO terms).
Were study selection criteria applied by two or more reviewers independently?	Yes	CS Appendix D.1.2 Both title and abstract screening and full text assessment were

		undertaken by two independent reviewers. Disagreements were resolved through discussion, or arbitration with a third independent reviewer when necessary.
Was data extraction performed by two or more reviewers independently?	Yes	CS Appendix D.1.2 Data extraction was performed by a single reviewer and checked by a second reviewer. Discrepancies between the reviewer and the person checking were resolved by a third independent reviewer
Was a risk of bias assessment or a quality assessment of the included studies undertaken? If so, which tool was used?	Yes	CS Appendix Table 9 Study quality was assessed using seven criteria. No reference is provided in the CS, but this appears to be the CRD assessment tool.[13]
Was risk of bias assessment (or other study quality assessment) conducted by two or more reviewers independently?	Unclear	The CS does not provide details of who performed the risk of bias assessment.
Is sufficient detail on the individual studies presented?	Partly	Yes for the semaglutide 2.4 mg trial (CS section B.2.3) but limited information given for the liraglutide 3.0 mg trial used in the indirect treatment comparison (ITC) (CS Appendix D.1.3.1). Some baseline characteristics were missing for the STEP 1 trial (clarification responses A6 & A11). Only aggregate baseline characteristics (pooled intervention and diet and physical activity arms) reported for liraglutide-eligible population in the ITC analysis (section 3.3.3.1).

If statistical evidence synthesis (e.g. pairwise meta-analysis, ITC, NMA) was undertaken, were appropriate methods used?	Yes	An ITC was undertaken, and we consider the methodology followed by the company is appropriate (see section 3.4).

3.2 Critique of studies of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

3.2.1 Included studies

The company's systematic literature review identified two potentially relevant phase 3 trials evaluating the efficacy of semaglutide 2.4 mg: STEP 1 and STEP 3 (CS section B.2.2). Both trials were conducted as part of the company's STEP clinical trial programme and were used to support the draft marketing authorisation. Both were sponsored by the company (Novo Nordisk). The company additionally provided information about 15 other ongoing or completed studies carried out as part of the STEP programme in CS section B.2.11, including reasons why the studies were excluded from the submission.

STEP 1 was an RCT comparing the efficacy of semaglutide 2.4 mg to placebo, both as adjuncts to a lifestyle intervention, in adults living with obesity (BMI ≥ 30 kg/m^2) or with overweight (BMI ≥ 27 kg/m^2) with at least one weight-related comorbidity. The trial did not include people with diabetes or HbA1c ≥ 6.5%. The company included STEP 1 in the CS review and use data from it in the economic model. The company provided the trial paper[5] and clinical study report[14] with the submission.

Throughout this report, we refer to semaglutide 2.4 mg in combination with the lifestyle intervention as 'semaglutide 2.4 mg' and placebo in combination with the lifestyle intervention as 'diet and physical activity'.

The design of the STEP 3 trial was the same as the STEP 1 trial, except that semaglutide 2.4 mg and placebo were given as adjuncts to intensive behavioural therapy (IBT). The trial was conducted solely in the United States.[3] The company state in CS section B.2.2 that IBT is not standard clinical practice in the UK, and, for this reason, they have excluded it from the CS review.

We have outlined what constituted IBT in the STEP 3 trial in Table 8, which also compares this intervention to standard management in the England and the lifestyle intervention used in the STEP 1 trial. In both the STEP 1 and STEP 3 trials, participants received individual counselling or IBT sessions. The ERG's clinical expert's advice indicates that neither of the interventions used in the STEP 1 and STEP 3 trials fully matches standard management in clinical practice. Clinical expert advice is that one-to-one counselling is not realistic in practice in England and people typically attend dietetics group sessions. We suggest the frequency of sessions offered in the STEP 1 trial more closely aligns to clinical practice in England than that in the STEP 3 trial. Clinical expert advice to the ERG is also that people's diet and activity levels, and therefore adherence, cannot be as closely monitored in practice as they were in the STEP 1 trial (in which participants recorded these daily in a diary or a smartphone application or other tools, which were then reviewed during counselling sessions).

We suggest that overall the standard management used in the STEP 1 trial more closely reflects practice in England than the IBT intervention used in STEP 3. Clinical expert advice to the ERG is that it is unlikely an NHS service could fund and provide the level of intervention delivered in the STEP 3 trial. However, whilst acknowledging this, we do not agree with the company's post-hoc exclusion of the STEP 3 trial from their systematic literature review. We believe the company should have included data from this trial in their submission. The trial met the inclusion criteria for the review and in our opinion, the comparator reflects the comparator specified in the NICE scope, in the sense that it was management of overweight and obesity without semaglutide that included a reduced calorie diet and increased physical activity. We suggest that standard management clinical practice is variable in England and so it is unlikely that an intervention used in a trial will fully reflect clinical practice. We provide selected results from the STEP 3 trial in section 3.6.

Table 8 Description of the standard management approaches used in clinical practice in England, the STEP 1 trial and the STEP 3 trial

Clinical practice [a]	STEP 1 [5]	STEP 3 [3]
People usually take part in one or two dietetics group sessions over 6 months (typically 6 to 9 sessions). They address healthy eating, having a balanced diet and eating behaviour,	Individual counselling sessions every 4 weeks during the 68-week intervention period of trial. The aim of these sessions was to help participants adhere to a reduced calorie	30 individual intensive behavioural therapy sessions with a dietician during the 68-week trial. The dietician gave the participants directions in physical activity, diet and

and include some behavioural intervention (motivational interviewing). One-to-one counselling is not realistic in practice, although some patients may receive one-to-one support for eating disorders. Some patients are also referred for physical activity intervention. Diet and physical activity are not recorded, so it is not possible to know how well people are adhering to this.	diet and increased physical activity. The aim of the diet element was to have a 500-kcal deficit per day compared to energy expenditure at baseline. Participants were encouraged to do 150 minutes of physical activity per week. Physical activity and diet were recorded daily and this record was reviewed during the counselling sessions.	behavioural strategies. Participants also had a hypo-caloric diet (1200-1800 kcal/d, depending on body weight at randomisation, after an initial 8-week low-calorie diet [1000-1200 kcal/d provided as meal replacements]) and were instructed to do 100 minutes of physical activity per week, titrated to 200 min/week during the trial.
[a] Our description of clinical practice here is based on information from our clinical expert about the form this typically takes.		

We have reviewed the other 15 ongoing or completed trials conducted as part of the STEP programme, which were outlined in CS section B.2.11. We agree with the company's exclusion of all them (the majority because they are ongoing or because the company stated data were not available in time for inclusion in the submission) except we are unclear whether or not the STEP 2 trial should have been included – we discuss this further in the next paragraph. As we discuss in section 3.2.1.3, we also note that the completed phase 3 trials STEP 5 and STEP 8 are relevant to the NICE scope and the decision problem, but the company stated data were not yet available for inclusion in the CS. The clinical trial reports for these studies are expected this year. We suggest data from these trials could potentially have a bearing on conclusions about the clinical effectiveness and cost-effectiveness of semaglutide 2.4 mg.

The STEP 2 trial evaluated the efficacy of semaglutide 2.4mg, 1.0 mg or placebo (all delivered alongside a lifestyle intervention, which involved a reduced-calories diet and increased physical activity) for weight management in people who were either overweight or obese (BMI ≥ 27 mg/m^2), had glycated haemoglobin 7-10% (53-86 mmol/mol) and who had been diagnosed with type 2 diabetes.[2] The STEP 2 trial semaglutide 2.4 mg and placebo arms meet the NICE scope and the decision problem. The company, however, state in CS Table 14 that the trial has not been included in the CS as the population – adults with type 2 diabetes – is not relevant to the submission. As outlined in section 2.3, in their clarification

response A1, the company explained that while semaglutide 2.4 mg might be used to treat weight in people with type 2 diabetes, clinical experts consulted by the company suggested that treatment for these patients would typically follow a diabetes treatment pathway where semaglutide would be used at a lower dose. We are unclear, having only spoken to one clinical expert, whether people with type 2 diabetes might be treated with the 2.4 mg dose in practice for the purposes of weight loss and maintenance. The company's clarification response and our clinical expert indicate this is possible. We suggest it is uncertain if the STEP 2 trial should have been included in the review, and further discussion with clinical experts during the appraisal process may help resolve this uncertainty.

We otherwise believe it is likely that all relevant studies of semaglutide 2.4 mg have been included in the CS (see section 3.2.1.3 for details about the ERG's additional searches for studies).

The trials identified for and included in the ITC are detailed in section 3.3.2.1.

3.2.1.1 Study characteristics

The company summarise the characteristics and methodology of the STEP 1 trial in CS section B.2.3.1. We have summarised the key characteristics of the trial in Table 14 and the outcomes assessed in Table 12 (in section 3.2.3 of this report), indicating which outcomes informed the CS economic model. The trial meets the decision problem and systematic literature review inclusion criteria. Semaglutide 2.4 mg was administered in line with the anticipated SmPC.

To be included in the trial, participants had to have one of the following weight-related comorbidities: hypertension, dyslipidaemia, obstructive sleep apnoea or cardiovascular disease. Clinical expert advice to the ERG is that these comorbidities are reflective of those seen in patients in practice who are likely to be treated with semaglutide 2.4 mg.

Table 9 STEP 1 trial characteristics

Trial characteristic	Description
Study design	Phase 3 double-bind, placebo-controlled RCT
Number and location of centres	129 sites in 16 countries, including 10 in the UK
Participant numbers	1,961 adults

Study population	Adults with obesity (BMI ≥ 30 kg/m^2), or overweight (BMI ≥ 27 kg/m^2) with at least one weight-related comorbidity, and without diabetes or HbA1c ≥ 6.5%
Comorbidities – eligibility criteria	To be included in the trial, participants with overweight had to have at least one of these weight-related co-morbidities (treated or untreated): hypertension, dyslipidaemia, obstructive sleep apnoea or cardiovascular disease
Intervention	Semaglutide 2.4 mg once weekly given as an adjunct to a lifestyle intervention[a]. Dose was titrated from a starting dose of 0.25 mg every four weeks to reach the maintenance dose.
Comparator	Matching placebo given as an adjunct to a lifestyle intervention [a].
Treatment and trial duration	Participants received semaglutide for 68 weeks, including 16 weeks of dose titration to reach the maintenance dose of 2.4 mg and a 52-week period of receiving the maintenance dose. A subset of participants then took part in a 52-week off-treatment extension phase where they did not receive semaglutide 2.4 mg or placebo nor the lifestyle intervention.
Stopping rule	A treatment non-responder stopping rule does not appear to have been used in the STEP 1 trial, but is applied in the CS economic model base case (see section 4.2.2).
Source: CS Table 3, CS Table 4, CS section B.2.3.1, STEP 1 trial paper,[5] CS Figure 3. [a] Details of the lifestyle intervention are given in our Table 8	

The company provide clinical efficacy results from the STEP 1 trial in the CS for the following population and subgroups:

- The whole trial population (full analysis set): people with a BMI ≥ 30 or ≥ 27 who have at least one of comorbidity (hypertension, dyslipidaemia, obstructive sleep apnoea [OSA] or cardiovascular disease [CVD])
- Subgroup: people with a BMI ≥ 30 plus at least one weight-related comorbidity
- Subgroup: people with a BMI ≥ 35 with non-diabetic hyperglycaemia and high CVD risk (this population matches the group of people for whom NICE recommends liraglutide 3.0 mg for the treatment of obesity in TA 664)[4]

3.2.1.2 Patients' baseline characteristics

The company present baseline characteristics for the STEP 1 trial full analysis set and the BMI ≥ 30 kg/m^2 plus ≥ one comorbidity subgroup in CS Table 5 and comment on these in CS

section B.2.3.2. In CS Table 5, race and BMI category characteristics were not reported for the subgroup, while they were provided for the full analysis set. The company provided this information in clarification response A6, attachment E. We have presented selected baseline characteristics in Table 10. Baseline characteristics for the BMI ≥ 35 kg/m^2 with non-diabetic hyperglycaemia and high CVD risk subgroup are provided in CS Table 11 and discussed in section 3.3.3.1 of this report.

We agree with the company that baseline characteristics were well balanced between the semaglutide 2.4 mg and diet and physical activity arms of the trial for both the BMI ≥ 30 kg/m^2 plus ≥ one comorbidity subgroup and full analysis set. We also agree with their conclusion that characteristics were similar across the full analysis set and the subgroup, with some expected higher rates of some disease characteristics in the subgroup, given their higher BMI.

The company state that clinicians considered the baseline characteristics of the trial reflected the UK obesity population, including the people who would typically be referred to SWMS. We understand from our clinical expert that in tier 3 services, people with higher BMIs than those in the STEP 1 trial are typically seen in practice and thus people have more comorbidities.

Table 10 Selected baseline characteristics of participants in the STEP 1 trial

	Trial population			
	BMI ≥ 30 mg/kg^2 plus ≥ 1 comorbidity (n = 1,470)		**Full analysis set (n = 1,961)**	
	Semaglutide 2.4 mg (n = 974)	**Diet and physical activity (n = 496)**	**Semaglutide 2.4 mg (n = 1,306)**	**Diet and physical activity (n = 655)**
Mean age, years (range)	[redacted]	[redacted]	46 (18–86)	47 (18–82)
Female, n (%)	696 (71.5)	375 (75.6)	955 (73.1)	498 (76.0)
Race, n (%)				
White	768 (78.9)	394 (79.4)	973 (74.5)	499 (76.2)
Asian	92 (9.4)	43 (8.7)	181 (13.9)	80 (12.2)
Black or African American	56 (5.7)	31 (6.3)	72 (5.5)	39 (6.0)
Other*	58 (6.0)	28 (5.6)	80 (6.1)	37 (5.6)
Hispanic or Latino ethnic group, n (%)	108 (11.1)	67 (13.5)	150 (11.5)	86 (13.1)
BMI				

	Trial population			
	BMI ≥ 30 mg/kg^2 plus ≥ 1 comorbidity (n = 1,470)		Full analysis set (n = 1,961)	
	Semaglutide 2.4 mg (n = 974)	Diet and physical activity (n = 496)	Semaglutide 2.4 mg (n = 1,306)	Diet and physical activity (n = 655)
Mean BMI, kg/m^2 (SD)	[redacted]	[redacted]	37.8 (6.7)	38.0 (6.5)
< 30 kg/m^2, n (%)	0	0	81 (6.2)	36 (5.5)
≥ 30 – < 35 kg/m^2, n (%)	319 (32.8)	158 (31.9)	436 (33.4)	207 (31.6)
≥ 35 – < 40 kg/m^2, n (%)	339 (34.8)	168 (33.9)	406 (31.1)	208 (31.8)
≥ 40 kg/m^2, n (%)	316 (32.4)	170 (34.3)	383 (29.3)	204 (31.1)
Patients with at least one comorbidity, n (%)	974 (100)	496 (100)	1048 (80.2)	532 (81.2)
Non-diabetic hyperglycaemia[a]	518 (53.2)	253 (51.0)	550 (42.1)	271 (41.4)

Source: this is a shortened version of CS Table 5, with additional information from the company's clarification response A6, attachment E.
[a] defined as haemoglobin A1c (HbA1c) levels in the range 6.0–6.4%, or fasting plasma glucose (FPG) levels in the range 5.5–6.9 mmol/L.

3.2.1.3 Ongoing studies

As discussed in section 3.2.1, the company provide a list of completed and ongoing studies on semaglutide 2.4 mg that are part of their STEP research programme. Among the studies listed are the completed phase 3 trials STEP 8 and STEP 5 (summarised in Table 11). The STEP 8 trial was a head-to-head comparison of semaglutide 2.4 mg with liraglutide 3.0 mg in people living with obesity or people with overweight who have at least one weight-related comorbidity. The company state that data from these trials were not available in time for this submission. Both the STEP 5 and STEP 8 trials are relevant to the decision problem for this appraisal, albeit it is unclear how many people in the STEP 8 trial might be included in a 'liraglutide-eligible' subgroup as per the NICE scope and decision problem.

Table 11 Details of the completed STEP 8 and STEP 5 trials

Trial (trial identifier), number of participants enrolled [a]	Population	Intervention and comparator(s)	Date clinical trial reports expected [b]
STEP 8 (NCT04074161) N = 338 participants	People living with obesity (BMI ≥ 30 kg/m^2) or people living with overweight (BMI ≥ 27 kg/m^2) with ≥ 1 weight-related comorbidity	• Semaglutide 2.4 mg • Liraglutide 3.0 mg • Placebo All administered during a 68-week treatment period and as an adjunct to a reduced-calorie diet and increased physical activity	Q4 2021
STEP 5 (NCT03693430) N = 304 participants	People living with obesity (BMI ≥ 30 kg/m^2) or people living with overweight (BMI ≥ 27 kg/m^2) with ≥ 1 weight-related comorbidity	• Semaglutide 2.4 mg • Placebo Both administered during a 104-week treatment period and as an adjunct to a reduced-calorie diet and increased physical activity	Q3 2021

Source: Table 14 and ClinicalTrials.gov

[a] The stated number of participants enrolled here is extracted from the number recorded under 'actual enrollment' on the ClinicalTrials.gov trial record.

[b] As stated in CS Table 14.

The company do not appear to have searched for other ongoing studies. For example, they have not searched trial registries. Given this gap in their searches, the ERG searched clinicaltrials.gov and the EU Clinical Trials Register for ongoing or recently completed trials of both semaglutide 2.4 mg and liraglutide 3.0 mg to check if any studies of either drug were missing from the submission and to check if there were any ongoing studies from which results may potentially be available soon. We did not identify any completed semaglutide 2.4 mg trials that had not been mentioned by the company in their submission or any additional

ongoing studies due to complete within the next 12 months. We did not identify any completed trials of liraglutide 3.0 mg or any that are due to complete within the next 12 months that could potentially inform the company's ITC.

ERG conclusion on included studies

The company have included one trial of semaglutide 2.4 mg in their systematic literature review; the STEP 1 trial. Baseline characteristics were well balanced between treatment arms. We understand from our clinical expert that in tier 3 services, people with higher BMIs than those in the STEP 1 trial are typically seen and thus people have more comorbidities. In this sense, we suggest the trial is not fully representative of the people who will likely be treated with semaglutide 2.4 mg in practice. We believe the company's exclusion of the STEP 3 trial was inappropriate and that data from the trial should have been included in the CS. It is uncertain whether or not semaglutide 2.4 mg will be used for weight loss and maintenance in people with type 2 diabetes in practice and therefore whether or not the STEP 2 trial should have been included in the submission. We consider the completed STEP 5 and STEP 8 trials are relevant to the appraisal (albeit it is unclear how many people in the STEP 8 trial might be included in a 'liraglutide-eligible' subgroup) but note that data from the trials are not yet available.

3.2.2 Risk of bias assessment

The company's quality (i.e. risk of bias) assessment for the STEP 1 trial is presented in CS Appendix D.3, based on Centre for Reviews and Dissemination (CRD) criteria.[13] The ERG assessed the STEP 1 trial using the same criteria, and the company's and ERG's judgements are provided in Appendix 1. The company and ERG conclude that STEP 1 was a well-conducted trial of good methodological quality and in general the ERG agree with the company's risk of bias judgements. However, the ERG are unclear about the risk of attrition bias in the company's analysis of STEP 1 (further details are provided in section 3.3.5).

In summary, the STEP 1 trial was generally well-conducted, but the ERG are unclear about the risk of attrition bias which introduces some uncertainty (of unknown magnitude and direction) to the outcome estimates reported in the CS.

3.2.3 Outcomes assessment

The efficacy outcomes assessed in the STEP 1 trial are summarised in CS Tables 3 and 4 and in Table 12 here. The company have included all the outcomes specified in the decision

problem and NICE scope in the CS, except for the incidence of type 2 diabetes (only reported at baseline) and cardiovascular events. Cardiovascular events do not appear to have been measured in the STEP 1 trial. The ITC report (section 3.2) states that there were few cases of type 2 diabetes to conduct statistical analyses.[15] In the economic model, longer-term benefit of weight loss on the incidence of diabetes and cardiovascular events is estimated using risk equations (CS section B.3.3.7). The company present changes in systolic blood pressure and fasting lipid profile from baseline as additional outcomes in the CS. These are not specified in the decision problem or the NICE scope, but are included in the CS as they inform the economic model. Efficacy results are presented as changes from baseline to week 68 (i.e. the end of the maintenance treatment period of the trial) or as status at week 68.

Table 12 Primary and other outcomes assessed in the STEP 1 trial

Outcome type	Outcomes assessed
Primary outcomes	• **Percentage change in body weight** from baseline to 68 weeks (the CS economic model uses the results of this outcome from the trial as efficacy inputs at months 4, 7 and 10, and years 1 and 2 in the economic model) Checking reviewer, please see CS section B.3.3.1.1. • Proportion of participants achieving a baseline body weight loss of ≥ 5% at 68 weeks
Other outcomes	• BMI (specifically, BMI change from baseline) • Weight loss (specifically: change in body weight in kg; and weight change ≥ 10%, ≥ 15% and ≥ 20%) • Incidence of type 2 diabetes (only reported at baseline) • Waist circumference • Glycaemic status (specifically: HbA1c (%) change from baseline; and, **percentage of participants with prediabetes or non-diabetic hyperglycaemia at baseline who achieved normoglycaemia** at 68 weeks) • Mortality • Adverse effects of treatment • Health-related quality of life • **Change in systolic blood pressure from baseline** • **Change in fasting lipid profile from baseline (specifically, HDL and total cholesterol)**
Source: CS Tables 3 and 4 Notes: Bold text shows the outcomes used in the economic model.	

The outcomes measured are appropriate and clinically relevant. Clinical expert advice to the ERG is that the key clinical outcomes for assessing the efficacy of treatment for obesity are weight loss, HbA1c and psychological and physical wellbeing. We suggest the latter would

be captured in the HRQoL outcomes included in the CS. One of the primary outcomes was the proportion of participants who achieved a ≥ 5% weight loss. This outcome is clinically meaningful. As referenced in the CS, a NICE clinical knowledge summary for the management of obesity[16] suggests a clinical aim of an overall reduction of 5-10% in body weight or higher in a person living with obesity. As we note in section 2.2.1, clinical expert advice to the ERG is that people typically achieve a weight loss of 5% in practice with the motivation of weight loss surgery and if they are able to engage with treatment.

The STEP 1 trial used the American Diabetes Association definition of prediabetes.[5] This defines prediabetes as an HbA1c level of 5.7 to 6.4% or FPG ≥ 5.6 mmol/L and ≤ 6.9 mmol/L, or two-hour post challenge (OGTT) FPG ≥ 7.8 mmol/L and ≤ 11.0 mmol/L. As outlined in CS section B.2.4.5, in the submission, the company have defined prediabetes in line with the definition of non-diabetic hyperglycaemia used in the NICE liraglutide appraisal (TA 664),[4] when presenting the achievement of normoglycaemia among participants who had non-diabetic hyperglycaemia at baseline in the STEP 1 trial for the FAS population, the target subgroup and the liraglutide-eligible subgroup. The CS states the TA 664 definition was an HbA1c level of 42 to 47 mmol/mol (6.0 to 6.4%) or a FPG level of 5.5 mmol/L. This is correct, but the upper bound FPG of 6.9 mmol/L[4] was missing from the definition in this section of the CS.

HRQoL was measured in the trial using the 36-Item Short Form Survey (SF-36) and the short form of Impact of Weight on Quality of Life-Lite for Clinical Trials (IWQOL-Lite-CT). The results of these measures were not used in the economic model, so we do not consider how these outcomes were measured further here. The model used published utility values (CS section B.3.4). We consider that the company's approach to estimating utility values is generally reasonable (see section 4.2.7).

ERG conclusion on outcomes assessment

The outcome measures included from the STEP 1 trial in the CS are appropriate and clinically relevant. No data are available from the trial on the longer-term outcomes of diabetes incidence and cardiovascular events. We have no concerns about how the outcomes were defined or measured.

3.2.4 Statistical methods of the included studies

3.2.4.1 Statistical procedures

The statistical procedures used in the STEP 1 trial are described in CS section B.2.4. The ERG have no concerns about the sample size calculation, the statistical approaches used for analysing each outcome or the methods used to impute missing data. The trial appears to be adequately powered.

3.2.4.2 Analysis sets

The company define the full analysis (FAS) and safety analysis sets in CS section B.2.4.1. The STEP 1 efficacy analysis used the FAS, which the company stated included all randomised participants in line with the intention-to-treat principal. The company defines two post-hoc subgroup analyses in section B.2.4.2.

The company provide clinical efficacy results in the submission for three trial populations:

- the whole trial population (FAS) (n = 1,961),
- the two post-hoc subgroups:
    - BMI ≥ 30 plus at least one weight-related comorbidity (n = 1,470), and
    - BMI ≥ 35 with non-diabetic hyperglycaemia and high CVD risk subgroup (n = 421).

The CS economic model base case, however, does not use the BMI ≥ 30 plus at least one weight-related comorbidity subgroup data and uses the FAS results instead for this population. The efficacy results from the subgroup are used in a company scenario analysis (see section 5.2.3).

3.2.4.3 Treatment estimands

Efficacy results are provided in the CS for two treatment estimands, shown in Table 13 and explained in CS section B.2.4.4. The company details what the term 'estimand' means in section B.1.2.4.4. Briefly, they are a way of handling intercurrent events that occur during a trial that might affect how the results are interpreted, such as a participant starting other medications (e.g. a rescue medication, a medication that the protocol prohibits or a subsequent therapy line).[17 18] A treatment policy estimand provides the treatment effect in the target population regardless of participants' discontinuation of the trial drug or use of other medications. The trial product estimand shows the treatment effect in the target population in the hypothetical situation that participants had continued using the trial medication and had not discontinued.[18] Therefore, the treatment policy estimand only

imputes data for participants who withdrew from the trial, while the trial product imputes data for participants using rescue medication, discontinuing the trial product, and withdrawing from the trial.

In the STEP 1 trial, the estimands were used to take into account the intercurrent events of participants starting other anti-obesity therapies (i.e. weight management drugs or weight loss surgery) and premature discontinuation.

Table 13 STEP 1 trial treatment estimands

Estimand (number of FAS participants included)	Definition
Treatment policy estimand (n = 1,961)	Estimated the effect of semaglutide 2.4 mg relative to diet and physical activity for all randomised participants regardless of starting other therapies, treatment adherence or premature discontinuation.
Hypothetical (trial product) estimand (n = 1,961)	Estimated the effect of semaglutide 2.4 mg relative to diet and physical activity for all randomised participants, assuming they remained on treatment and did not start other anti-obesity therapies (i.e. this estimand excludes the effects of other anti-obesity therapies and any effects after treatment discontinuation)
Source: CS section B.2.4.4.	

The treatment policy estimand was used for regulatory approval. We believe, and clinical advice to the ERG suggests, that the treatment policy estimand results are the most relevant to clinical practice. The trial product estimand was used in the economic model alongside a treatment stopping rule. The effects of other anti-obesity therapies are estimated in the model using published literature. The ERG considers the use of the trial product estimand to incorporate the effect of treatment discontinuation to be a reasonable and appropriate approach (section 4.2.6.1). The company have conducted a scenario analysis with no stopping rule applied, which uses the treatment policy estimand (section 5.2.3).

The company do not compare baseline characteristics between the participants included in each of these estimands, so it is not possible to determine how the participants whose data generated the trial product estimands differed from or were similar to participants in the treatment policy estimand.

ERG comment on study statistical methods

We have not identified any issues with the statistical methods of the STEP 1 trial.

3.2.5 Efficacy results of the intervention studies

Here we provide the results of the outcomes from the STEP 1 trial that inform the economic model, namely:

- Percentage change in body weight
- Percentage of participants with prediabetes or non-diabetic hyperglycaemia at baseline who achieved normoglycaemia
- Change in systolic blood pressure
- Change in HDL and total cholesterol

See Appendix 2 of this report for the results of the following other outcomes measured in the STEP 1 trial: other weight loss outcomes, percentage of participants with a specified weight change from baseline, waist circumference change, incidence of type 2 diabetes (only reported at baseline and as a safety outcome), HbA_{1c} (%) change from baseline, and HRQoL.

3.2.5.1 Percentage change in weight from baseline at 68 weeks

Across the three populations and two estimands analysed, the percentage decrease in weight from baseline to 68 weeks ranged from 14.2 to 16.9 percentage points in the semaglutide 2.4 mg arm, and from 2.41 to 2.82 percentage points in the diet and physical activity arm (Table 14). The difference between trial arms was statistically significant for the FAS population (95% CIs exclude zero) but confidence intervals were not reported for the other analysis populations.

Table 14 Percentage change in weight from baseline at 68 weeks

Estimand (Data source)	Semaglutide 2.4 mg		Diet and physical activity		Difference (95% CI)
	Mean (SD[a]) change	N	Mean (SD[a]) change	N	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Treatment policy (CSR 14.2.9)	-14.85 %-points	1306	-2.41 %-points	655	-12.44% (-13.37 to -11.51); p<0.0001

Trial product (CSR 14.2.20)	-16.86 %-points	1306	-2.44 %-points	655	-14.42 (-15.29 to -13.55); p<0.0001
Target subgroup (BMI ≥30 plus ≥1 weight-related comorbidity) (post hoc analysis)					
Treatment policy (CS B.2.7.1)	-14.8 %-points	974	-2.6 %-points	496	-12.2 [b]
Trial product (Appendix E.2)	-16.59 (8.85) %-points	974	-2.56 (8.99) %-points	496	-14.03 [b]
Liraglutide-eligible subgroup (BMI≥35 with non-diabetic hyperglycaemia and CVD risk) (post hoc analysis)					
Treatment policy (CS B.2.7.2)	-14.2 %-points	273	-2.8 %-points	148	-11.4 [b]
Trial product (Appendix E.2)	-15.89 (8.87) %-points	273	-2.82 (9.00) %-points	148	-13.07 [b]
FAS: full analysis set [a] SD reported for some analyses [b] Not reported; raw difference calculated by reviewer					

3.2.6 Glycaemic status

The proportion of patients with prediabetes or non-diabetic hyperglycaemia at baseline who achieved normoglycaemia at week 68 was clearly higher for the semaglutide 2.4 mg arm than the diet and physical activity arm (Table 15). NB this outcome was not reported for the trial product estimand. The proportion who achieved normoglycaemia informs the economic model (CS section B.3.3.1.2) although there is a discrepancy between the data used in the model (CS Table 22) and those reported from the STEP 1 trial.

Table 15 Percentage of participants with prediabetes or non-diabetic hyperglycaemia at baseline who achieved normoglycaemia at 68 weeks (treatment policy estimand)

Baseline population	Semaglutide 2.4 mg		Diet and physical activity		Difference
	%	N [a]	%	N [a]	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Participants shifting from prediabetes to normo-glycaemic (trial publication)	84.1%	593	47.8%	263	36.3 [b]

Participants shifting from non-diabetic hyperglycaemia to normo-glycaemic (CS Table 10)	79.8%	550	39.1%	271	40.7 [b]
Target subgroup (BMI ≥30 plus ≥1 weight-related comorbidity) (post hoc analysis)					
Participants shifting from non-diabetic hyperglycaemia to normo-glycaemic	79.2%	518	20.0%	253	59.2 [b]
Liraglutide-eligible subgroup (BMI≥35 with non-diabetic hyperglycaemia and CVD risk) (post hoc analysis)					
Participants shifting from non-diabetic hyperglycaemia to normo-glycaemic (CS Table 10)	78.4%	273	36.5%	148	41.9 [b]

FAS: full analysis set

[a] The denominator is the number of patients with prediabetes or non-diabetic hyperglycaemia at baseline

[b] Not reported; raw difference calculated by reviewer

3.2.6.1 Systolic blood pressure

Across the analyses conducted, mean systolic blood pressure decreased from baseline to week 68 by 6.2 to 8.6 mmHg in the semaglutide 2.4 mg arm and by 1.0 to 2.2 mmHg in the diet and physical activity arm (Table 16). The difference between trial arms was statistically significant for the FAS population (95% CIs exclude zero) but confidence intervals were not reported for the other analysis populations. The company do not comment on the clinical significance of these changes in systolic blood pressure, which we note are relatively small.

Table 16 Systolic blood pressure change from baseline at 68 weeks

Estimand (Data source)	**Semaglutide 2.4 mg**		**Diet and physical activity**		**Difference (95% CI)**
	Mean (SD[a]) change	**N**	**Mean (SD[a]) change**	**N**	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Treatment policy (CSR 14.2.87)	-6.16 mmHg	1306	-1.06 mmHg	655	-5.10 (-6.34 to -3.87); p<0.0001

Trial product (CSR 14.2.150)	-7.08	1306	-1.14	655	-5.93 (-7.19 to -4.68); p<0.0001
Target subgroup (BMI ≥30 plus ≥1 weight-related comorbidity) (post hoc analysis)					
Treatment policy (CS B.2.7.1)	-6.4 mmHg	974	-1.0 mmHg	496	-5.4 [b]
Trial product (Appendix E.2)	-7.25 (13.08) mmHg	974	-1.39 (13.50) mmHg	496	-5.86 [b]
Liraglutide-eligible subgroup (BMI≥35 with non-diabetic hyperglycaemia and CVD risk) (post hoc analysis)					
Treatment policy (CS B.2.7.2)	-7.7 mmHg	273	-1.6 mmHg	148	-6.1 [b]
Trial product (Appendix E.2)	-8.55 (13.06)	273	-2.23 (13.27)	148	-6.32 [b]
FAS: full analysis set [a] SD reported for some analyses [b] Not reported; raw difference calculated by reviewer					

3.2.6.2 Fasting HDL and total cholesterol

HDL cholesterol

The company do not consistently report the change in fasting HDL cholesterol from baseline to 68 weeks for all the subgroups and estimands analysed. Where reported, the data suggest that in the FAS population HDL cholesterol increased marginally from baseline up to week 68, slightly more so in the semaglutide 2.4 mg arm (Table 17). But the changes appear very small (<0.5 mg/dL, with ratios close to 1.0). The company do not comment on the clinical significance of these changes in HDL cholesterol, although they report that the difference between trial arms is statistically significant.

Table 17 Geometric mean fasting HDL cholesterol (mg/dL) ratio to baseline and mean change at 68 weeks

Estimand (Data source)	**Semaglutide 2.4 mg**		**Diet and physical activity**		**Ratio difference (95% CI)**
	Ratio (mean [SD] change [a])	**N**	**Ratio (mean [SD] change [a])**	**N**	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					

Treatment policy (CSR 14.2.96) (CS B.2.6.5)	1.05 (0.04 mg/dL [a])	1306	1.01 (0.01 mg/dL [a])	655	1.04 (1.02 to 1.05); p<0.0001
Trial product (CSR 14.2.151; Appendix E.2)	1.05 (0.05 [0.16] mg/dL)	1306	1.02 (0.02 [0.17] mg/dL)	655	1.03 (1.02 to 1.05); p<0.0001
Target subgroup (BMI ≥30 plus ≥1 weight-related comorbidity) (post hoc analysis)					
Treatment policy (CS B.2.7.1)	1.0 [b] (0.0 mg/dL)	974	1.0 [b] (0.00 mg/dL)	496	1.0 [b]
Trial product (Appendix E.2)	NR (0.05 [0.16] mg/dL)	974	NR (0.02 [0.17] mg/dL)	496	NR
Liraglutide-eligible subgroup (BMI≥35 with non-diabetic hyperglycaemia and CVD risk) (post hoc analysis)					
Treatment policy (CS.B.2.7.2)	NR (0.1 mg/dL)	273	NR (0.0 mg/dL)	148	NR
Trial product (Appendix E.2)	NR (0.08 [0.16] mg/dL)	273	NR (0.02 [0.16] mg/dL)	148	NR

FAS: full analysis set; NR: not reported

[a] log scale; SD reported for some analyses

[b] Not reported; calculated by reviewer

Total cholesterol

The company do not consistently report the change in fasting total cholesterol from baseline to 68 weeks for all the subgroups and estimands analysed. Where reported, the data across the analyses conducted suggest that in the FAS population total cholesterol decreased marginally or remained stable from baseline up to week 68, changing by 0 to -0.04 mg/dL in the semaglutide 2.4 mg arm and with no change in the diet and physical activity arm (Table 58). The company do not comment on the clinical significance of these changes in HDL

cholesterol, although they report that the difference between trial arms is statistically significant.

Table 18 Geometric mean fasting total cholesterol (mg/dL) ratio to baseline and mean change at 68 weeks

Estimand (Data source)	Semaglutide 2.4 mg		Diet and physical activity		Ratio difference (95% CI)
	Ratio (mean [SD] change [a])	N	Ratio (mean [SD] change [a])	N	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Treatment policy (CSR 14.2.96) (CS B.2.6.5)	0.97 (-0.04 mg/dL)	1306	1.00 (0.00 mg/dL)	655	0.97 (0.95 to 0.98); p<0.0001
Trial product (CSR 14.2.151; Appendix E.2)	0.96 (-0.04 [0.16] mg/dL)	1306	1.00 (0.00 [0.16] mg/dL)	655	0.96 (0.94 to 0.97)
Target subgroup (BMI ≥30 plus ≥1 weight-related comorbidity) (post hoc analysis)					
Treatment policy (CS B.2.7.1)	1.0 [c] (-0.0 mg/dL)	974	1.0 [c] (0.0 mg/dL)	496	1.0 [b]
Trial product (Appendix E.2)	NR (-0.04 [0.16] mg/dL)	974	NR (0.00 [0.16] mg/dL)	496	NR
Liraglutide-eligible subgroup (BMI≥35 with non-diabetic hyperglycaemia and CVD risk) (post hoc analysis)					
Treatment policy (CS.B.2.7.2)	1.0 [c] (0.0 mg/dL)	273	1.0 [c] (0.0 mg/dL)	148	1.0 [b]
Trial product (Appendix E.2)	NR (-0.04 [0.17] mg/dL)	273	NR (-0.02 [0.17] mg/dL)	148	NR

FAS: full analysis set; NR: not reported

[a] log scale; SD reported for some analyses

[b] Not reported; calculated by reviewer

3.2.6.3 Subgroup analyses

The STEP 1 trial results for the target and liraglutide-eligible subgroups have been reported alongside those for the full analysis set population above, to make it easier for the reader to make comparisons between the groups.

3.2.7 Safety outcomes

The majority of participants (>85%) in both the semaglutide 2.4 mg arm and the diet and physical activity (plus placebo) arm of the STEP 1 trial experienced adverse events. The rate of any adverse events was marginally more frequent in the semaglutide 2.4 mg arm than the diet and physical activity arm (89.7% versus 86.4%), as was the rate of serious adverse events (9.8% versus 6.4%). Overall, the rate of adverse events per 100 person-years was higher in the semaglutide 2.4 mg arm (566.1) than the diet and physical activity arm (398.0) (Table 19).

Adverse events led to discontinuations in 7.0% of those receiving semaglutide 2.4 mg and 3.1% of those receiving diet and physical activity, with discontinuations due to gastrointestinal disorders being the main adverse event leading to discontinuation.

One death was reported in each trial arm, neither of which was considered by the independent external event adjudication committee to be related to semaglutide 2.4 mg or diet and physical activity (CS Appendix F.1).

The rate of adverse events considered probably related to treatment was relatively high for the diet and physical activity arm, i.e. for participants receiving placebo and the lifestyle intervention (22.4%).

The CS reports the most frequent adverse events, i.e. those which affected ≥10% of participants in either trial arm (Table 19) but does not specify the rates of grade 3 or grade 4 events. Rates of nasopharyngitis and upper respiratory tract infection did not differ between the trial arms whereas the other common adverse events, which were mostly gastrointestinal disorders, were more frequent in the semaglutide 2.4 mg arm.

Table 19 Summary of adverse events

Adverse event (AE)	Semaglutide 2.4 mg		Diet and physical activity	
	Participants N (%)	Events per 100 person-years	Participants N (%)	Events per 100 person-years
Any AE	1171 (89.7)	566.1	566 (86.4)	398.0
Serious AE	128 (9.8)	9.6	42 (6.4)	6.4
AE leading to discontinuation	92 (7.0)	7.2	20 (3.1)	2.8
GI disorders leading to discontinuation	59 (4.5)	4.6	5 (0.8)	0.6
Mortality	1 (0.1)	0.1	1 (0.2)	0.3
Treatment-related AE				
Probably related	571 (43.7)	125.9	147 (22.4)	39.8
Possibly related	726 (55.6)	158.3	223 (34.0)	66.9
AE reported in ≥10% of participants in either trial arm				
Nausea	577 (44.2)	62.6	114 (17.4)	17.6
Diarrhoea	412 (31.5)	44.9	104 (15.9)	16.6
Vomiting	324 (24.8)	37.3	43 (6.6)	6.3
Constipation	306 (23.4)	22.9	62 (9.5)	8.8
Nasopharyngitis	281 (21.5)	28.1	133 (20.3)	26.0
Headache	198 (15.2)	22.7	80 (12.2)	12.5
Dyspepsia	135 (10.3)	10.5	23 (3.5)	3.6
Abdominal pain	130 (10.0)	10.3	36 (5.5)	4.9
Upper RT infection	114 (8.7)	9.3	80 (12.2)	14.0

GI: gastrointestinal; RT: respiratory tract Source: CS Table 13 and CS Appendix F.2

The company report a set of adverse events which they refer to as being of "particular interest" (CS Appendix F.2) or "safety focus areas" (trial publication), which were selected "based on therapeutic experience with glucagon-like peptide-1 receptor agonists and in line with regulatory feedback and requirements" (CS section B.2.10.2). The most frequent events of particular interest were gastrointestinal disorders, which occurred in 74.2% of participants in the semaglutide 2.4 mg arm and 47.9% of participants in the diet and physical activity arm (Table 20). Cardiovascular events, which are specified as an outcome in the Decision Problem (CS section B.1.1) and inform the economic analysis (section B.3.3.71), are

included among the events of particular interest. However, cardiovascular events are only presented at an aggregate level for each arm of the STEP 1 trial and are not defined explicitly.

Table 20 Safety focus areas

Adverse event	Semaglutide 2.4 mg		Diet and physical activity	
	Participants N (%)	Events per 100 person-years	Participants N (%)	Events per 100 person-years
GI disorders	969 (74.2)	252.6	314 (47.9)	89.1
Gallbladder-related	34 (2.6)	2.5	8 (1.2)	1.0
> Hepatobiliary	33 (2.5)	2.3	5 (0.8)	0.6
>> Cholelithiasis	23 (1.8)	1.4	4 (0.6)	0.5
Hepatic disorders	31 (2.4)	2.2	20 (3.1)	2.9
Acute pancreatitis	3 (0.2)	0.2	0	0
Cardiovascular disorders	107 (8.2)	7.2	75 (11.5)	10.5
Allergic reactions	96 (7.4)	6.3	54 (8.2)	7.6
Injection site reactions	65 (5.0)	5.8	44 (6.7)	9.9
Malignant neoplasms	14 (1.1)	0.8	7 (1.1)	0.8
Psychiatric disorders	124 (9.5)	9.4	83 (12.7)	13.6
Acute renal failure	3 (0.2)	0.2	2 (0.3)	0.2
Hypoglycaemia	8 (0.6)	0.9	5 (0.8)	0.8

GI: gastrointestinal Source: CS Appendix F.2 and trial publication

3.2.8 Pairwise meta-analysis of intervention studies

As the company only included one trial (STEP 1) comparing semaglutide 2.4 mg to diet and physical activity, the company did not undertake a meta-analysis.

3.3 Critique of studies included in the indirect comparison and/or multiple treatment comparison

The company conducted an indirect treatment comparison (ITC) to compare semaglutide 2.4 mg/week against liraglutide 3.0 mg/day using the placebo plus diet and physical activity arms of the STEP 1 and SCALE 1839 trials as the common comparator. The ITC utilised

data from the liraglutide-eligible subgroup of patients, i.e. those with BMI≥35 kg/m^2 with non-diabetic hyperglycaemia and high CVD risk (clarification response A12).

3.3.1 Rationale for the ITC

A direct comparison of semaglutide 2.4mg vs liraglutide 3.0mg is being conducted in the recently completed STEP 8 trial; however, STEP 8 results will not be available until Q4 2021. In the absence of any direct comparisons, an indirect comparison was deemed appropriate by both the company (CS section B.2.9) and the ERG. We assume that when results of the STEP 8 trial become available they would supersede the results of the ITC, i.e. the role of the ITC is for interim decision making.

3.3.2 Identification, selection and feasibility assessment of studies for the ITC

3.3.2.1 Identification of studies

A SLR was conducted to identify relevant studies for inclusion in the indirect comparisons of semaglutide 2.4 mg versus liraglutide 3.0 mg (CS section B.2.9). Details of the SLR and a summary of the included studies are provided in CS Appendix D.1.2, CS Appendix D.1.3 and an ITC Report.[15] The search was conducted in September 2020 and updated in April 2021 and was considered by the ERG to be broadly up to date and fit for purpose. We consider it is unlikely that key studies have been missed.

3.3.2.2 Selection of studies

Study selection is reported in CS Appendix D.1.2. The company's eligibility criteria (summarised in Table 21) are generally broader than the decision problem. Apart from HRQoL, all outcomes specified in the decision problem are captured in the eligibility criteria. Broad reasons for excluding studies at full text screening are provided in CS Appendix Table 5. The ERG requested further details of the excluded studies in order to check whether the company's exclusions were appropriate (clarification response A2).

Table 21 Eligibility criteria for the indirect treatment comparison

PICOD criterion	Inclusion criteria
Population	Adults with: • BMI ≥ 27 kg/m^2 and one weight-related co-morbidity • BMI ≥ 30 kg/m^2 (with weight-related co-morbidities) NB CS Appendix Table 4 states people with "BMI ≥ 30 kg/m^2 (without weight-related co-morbidities)" were included; however according to

	clarification response A3 people without weight-related comorbidities were excluded.
Intervention	Semaglutide 2.4 mg
Comparators	**As per the decision problem:** • Standard management without semaglutide (including a reduced calorie diet and increased physical activity) • Liraglutide, 3.0 mg (Saxenda)
Outcomes	**Outcomes consistent with the decision problem:** • Proportion of subjects losing at least 5%, 10%, and 15% of baseline fasting body weight • Weight loss in kg • Mean % change in weight • HbA1c - Mean % change in HbA1c versus baseline HbA1c • % reversing from prediabetes to normal glucose tolerance • Waist circumference • Safety outcomes (incidence of hypoglycaemia, incidence of SAEs and discontinuations due to AEs) **Outcomes additional to the decision problem:** • Systolic blood pressure (SBP) - Absolute change in mm Hg vs baseline • HDL – Absolute change in mg/dl versus baseline • Total cholesterol - Absolute change in mg/dl versus baseline • % reduction in antihypertensive treatment • % change in glucose lowering drugs **Outcomes stated in the decision problem but not included:** • HRQoL
Design	RCTs with data following >9 months of treatment duration
Source: summary of CS Appendix Table 4. This table gives a shortened overview of the key criteria and is not exhaustive (i.e. the language criteria are not included here).	

The company's process for eligibility screening followed good practice, with titles, abstracts and full-text articles assessed by two reviewers independently.

3.3.2.3 Studies eligible for inclusion in the ITC

Following the study selection process, the company identified 3 relevant RCTs (reported in 7 references):

- STEP 1 and STEP 3 trials of semaglutide 2.4 mg versus placebo, both used as adjuncts to diet and physical activity in overweight or obese patients;
- The SCALE 1839 Obesity and Prediabetes study of liraglutide 3.0 mg/day versus placebo, both used as adjuncts to diet and physical activity in overweight or obese patients. For simplicity, we refer to liraglutide 3.0 mg/day in combination with diet and physical activity as 'liraglutide 3.0 mg' throughout this report.

The company excluded the STEP 3 trial post-hoc because all enrolled patients received IBT in addition to their randomised treatment (i.e. placebo or semaglutide 2.4 mg), which the company argue is not considered standard practice in the UK (CS Appendix section D.1.3). As explained in section 3.2.1 above the ERG disagree with the company and believe the STEP 3 trial meets the NICE scope and decision problem and should have been included in the company's analyses.

3.3.2.4 Included studies and populations

Following the selection process outlined above the ITCs included data from two trials: STEP 1, providing patient data for semaglutide 2.4 mg; and SCALE 1839 obesity and pre-diabetes, providing patient data for liraglutide 3.0 mg. Both trials were conducted by the company, and the company used individual participant data (IPD) for the ITC analyses.

Dose escalation and study duration for patients with prediabetes

In STEP 1 the target dose (2.4mg semaglutide) was achieved after 16 weeks and end of treatment was at 68 weeks, i.e. after 52 weeks on the target dose. In SCALE 1839 the target dose (3.0mg liraglutide) was achieved after 4 weeks and end of treatment was at 56 weeks, i.e. after 52 weeks on the target dose (CS Appendix D.1.3.1). The company conducted a base case comparison based on time after randomization (approximately 1 year) rather than weeks on the target dose of treatment: the analysis was conducted for 52 weeks after randomization in STEP 1 and 56 weeks after randomization in SCALE 1839 (ITC Report), with scenario analyses at other time points. The company do not provide a rationale for this approach, which does not reflect the full treatment duration in STEP 1 (36 weeks of the target dose rather than 52 weeks, and hence the company regard this as a conservative comparison; ITC Report). The ERG note that the benefit of semaglutide 2.4 mg appears to be established before the full 68-week trial duration and therefore the company's analysis approach appears reasonable and conservative.

Subgroups analysed

The company's main population of interest, referred to as the ITC base case, was the sub-population of participants in each trial who had BMI≥35 kg/m^2, non-diabetic hyperglycaemia (high risk of` diabetes – NICE definition), and high risk of cardiovascular disease (CVD). The ITC report states that in addition to the base case, (unadjusted) analyses were also conducted for the "broader subpopulation of patients with pre-diabetes" (specific BMI range not stated) (these analyses are not relevant to this appraisal).

Definitions

In both trials **prediabetes** was defined according to American Diabetes Association criteria: HbA1c 5.7−6.4% both inclusive or 5.5 mmol/L ≤FPG ≤6.9 mmol/L or 2-hour post-challenge (oral glucose tolerance test [OGTT]) plasma glucose ≥7.8 mmol/L and ≤11.0 mmol/L. **Normoglycaemia** is not explicitly defined in the ITC Report,[15] but the ITC report states **high risk of diabetes** is defined by NICE as having 5.5 mmol/L ≤FPG ≤6.9 mmol/L or 6.0% ≤HbA1c ≤6.4% (this definition was used in the liraglutide appraisal [TA664]). The ITC report states the definition of **high risk of CVD** in the liraglutide appraisal was total cholesterol >5 mmol/L or systolic blood pressure >140 mmHg or HDL <1.0 mmol/L for men and <1.3 mmol/L for women, and this is correct.

3.3.3 Clinical heterogeneity assessment

3.3.3.1 Trial baseline characteristics

The CS presents baseline characteristics for STEP 1 (CS Tables 5 and 6; CS Appendix Table 10) and SCALE 1839 (CS Table 11). However, comparisons of baseline characteristics across both arms in both trials is only possible for the FAS population (Table 22). An aggregate comparison (intervention and placebo arms pooled) between STEP 1 and SCALE 1839 for the liraglutide-eligible subgroup is provided in CS Table 11, reproduced in Table 23 below.

Overall, the STEP 1 and SCALE 1839 trial populations were similar, although STEP 1 had a slightly higher proportion of Asian participants and lower proportion of people with prediabetes (Table 22). Fewer baseline characteristics are reported for the liraglutide-eligible subgroup, which is the primary population of interest for the ITC (Table 23), as the company have only presented variables which they believe are potential effect modifiers. Where reported, the baseline characteristics of the liraglutide-eligible subgroup of participants were also generally similar between STEP 1 and SCALE 1839; the largest differences were in

white ethnicity (6.2%-points higher in SCALE 1839), dyslipidaemia (5%-points higher in STEP 1), Asian ethnicity (4.5%-points higher in STEP 1) and mean weight (1.3 kg higher in STEP 1).

Table 22 Baseline characteristics of STEP 1 and SCALE 1839 trials: FAS populations

Variable **Mean (SD) unless stated otherwise**		**STEP 1**		**SCALE 1839**	
		Semaglutide 2.4 mg N=1306	**Diet and physical activity N=655**	**Liraglutide 3.0 mg N=2487**	**Diet and physical activity N=1244**
Age, years		46 (13)	47 (12)	45.2 (12.1)	45.0 (12.0)
Female sex, %		73.1	76.0	78.7	78.1
Self-reported race or ethnic group, %	White	74.5	76.2	84.7	85.3
	Black [a]	5.5	6.0	9.7	9.2
	Asian	13.9	12.2	3.6	3.7
	Other	6.1	5.6	1.9	1.8
Body weight, kg		105.4 (22.1)	105.2 (21.5)	106.2 (21.2)	106.2 (21.7)
BMI, kg/m^2		37.8 (6.7)	38.0 (6.5)	38.3 (6.4)	38.3 (6.3)
Waist circumference, cm		114.6 (14.8)	114.8 (14.4)	115.0 (14.4)	114.5 (14.3)
HbA1c %		5.7 (0.3)	5.7 (0.3)	5.6 (0.4)	5.6 (0.4)
Prediabetes, %		45.4	40.2	61.4	60.9
Systolic BP, mmHg		126 (14)	127 (14)	123.0 (12.9)	123.2 (12.8)
Total cholesterol, mg/dL (% CV [b])		189.6 (20.5) [b]	192.1 (19.4) [b]	193.7 (19.1) [b]	194.3 (18.8) [b]
HDL cholesterol, mg/dL (% CV [b])		49.4 (25.6) [b]	49.5 (25.0) [b]	51.4 (26.2) [b]	51.0 (26.4) [b]
Dyslipidaemia, %		38.2	34.5	29.6	28.9
Hypertension, %		36.1	35.7	34.2	35.9
On anti-hypertensive drug, %		23.8	23.2	NR [c]	NR [c]
On lipid-lowering drug, %		19.1	17.4	NR [c]	NR [c]

Sources: STEP 1: trial publication[5] and CS Table 5; SCALE 1839: trial publication[19]
BP: blood pressure; CV: coefficient of variation; NR: not reported
[a] reported as Black or African American in the STEP 1 trial
[b] geometric mean and % coefficient of variation
[c] reported for prediabetic and normoglycaemia groups but not FAS population

Table 23 Baseline characteristics of STEP 1 and SCALE 1839 trials: liraglutide-eligible populations (BMI≥35, non-diabetic hyperglycaemia and high cardiovascular risk)

Variable		STEP 1 N=421 [a]	SCALE 1839 N=800 [a]
Age, years, mean (SD)		48.1 (12.06)	48.2 (11.24)
Female, n/N (%) [b]		314/421 (74.6)	606/800 (75.8)
Race / ethnicity, n/N (%) [c]	White	334/421 (79.3)	684/800 (85.5)
	Black or African American	23/421 (5.5)	74/800 (9.3)
	Asian	34/421 (8.1)	29/800 (3.6)
	Other	18/421 (4.3)	13/800 (1.6)
	Not reported	12/421 (2.9)	0/800 (0)
Weight, kg, mean (SD) [b]		117.2 (21.91)	115.9 (19.76)
BMI, kg/m^2, mean (SD)		42.1 (6.28)	41.7 (5.35)
Waist circumference		Not reported	Not reported
HbA_{1c}, %, mean (SD) [b]		5.9 (0.28)	5.8 (0.34)
Systolic BP		Not reported	Not reported
Total cholesterol		Not reported	Not reported
HDL cholesterol		Not reported	Not reported
CVD, n/N (%)		36/421 (8.6)	88/800 (11.0)
Dyslipidaemia, n/N (%)		164/421 (39.0)	272/800 (34.0)
Hypertension, n/N (%)		190/421 (45.1)	389/800 (48.6)

Source: reproduction of CS Table 11 with minor modification
[a] The sample sizes given in CS Table 11 are for the FAS populations. The correct subgroup sample sizes were confirmed by the company in clarification response A9
[b] The CS states that these variables were considered potential effect modifiers and included in adjustment 1; age, dyslipidaemia, hypertension and cardiovascular disease were additionally included in adjustment 2.
[c] From clarification response A11 (not reported in the CS)

Overall the trials appear generally well-balanced in terms of the key prognostic variables that are relevant in obesity management.

3.3.3.2 Effect modifiers

3.3.3.2.1 Potential effect modifiers of drug exposure

The company explored the factors which affect exposure to semaglutide and liraglutide (ITC Report section 2.4.1). They considered exposure to semaglutide up to 1.0 mg/week in a diabetic population in a study by Carlsson Petri et al.[20] (ITC Report Figure 4) and exposure to liraglutide up to 3.0 mg/day in a population with obesity, in a study by Overgaard et al.[21] (ITC Report Figure 3). The company do not comment on whether other data sources were available or whether the factors affecting semaglutide exposure to a maximum of 1.0

mg/week would also apply to the intended 2.4 mg/week dose. As reported in the literature,[20] [21] the company conclude that baseline body weight (inversely related to exposure) and, for liraglutide only, sex (lower exposure in men) were the only effect modifiers for drug exposure (the reason why sex should be an effect modifier for liraglutide but not semaglutide is not discussed). For both drugs there were statistically significant effects on exposure of age, race, ethnicity, baseline glycaemic status, injection site and renal function (ITC Report Figures 3 and 4), but the company state race, ethnicity and age were not found to have a clinically relevant effect on exposure, which is consistent with the conclusions of the cited studies[20] [21] (the studies also reported no clinically relevant effects of sex, age, race, ethnicity, renal function, or injection site on exposure to semaglutide 1.0 mg;[20] and no clinically relevant effects of age ≥70 years, race, ethnicity and glycaemic status on exposure to liraglutide 3.0 mg[21]). As noted above, we believe there is some uncertainty in how generalisable these findings are beyond the specific populations and drug dosing in these studies.

3.3.3.2.2 Potential effect modifiers of relative weight change

The company identified baseline body weight/BMI and gender as potential effect modifiers of relative change in body weight based on subgroup analyses for semaglutide and liraglutide respectively versus placebo (ITC Report section 2.4.2.2). However, these subgroup analyses are not presented.

3.3.3.2.3 Potential effect modifiers of waist circumference, systolic blood pressure and lipids

The company argue that the treatment effect of liraglutide versus placebo on waist circumference, systolic blood pressure and lipids was predominantly impacted by the treatment effect on relative weight loss (ITC Report section 2.4.2.3). Accordingly, the effect modifiers for waist circumference, systolic blood pressure and lipids would be the same as those for weight loss. This observation is based on analysis of data from a series of SCALE trials by Bays et al.[22] (ITC Report Figure 5).

3.3.3.2.4 Potential effect modifiers of HbA1c and glycaemic status

The company cite evidence that the treatment effect of liraglutide versus placebo depends on baseline HbA_{1c} in diabetic populations[23] and they argue that the exposure of GLP-1 RAs is not expected to differ between diabetic and non-diabetic populations. The company's conclusion is that, in addition to gender and weight, baseline HbA_{1c} is a relevant effect modifier to consider in the ITC (ITC Report section 2.4.2.4).

As shown in Table 23 above, the effect modifiers weight, sex and HbA_{1c} were similar for the liraglutide-eligible subgroup in the STEP 1 and SCALE 1839 trials, apart from a slight difference in mean weight (1.3 kg higher in STEP 1).

3.3.4 Similarity of treatment effects

The ITC uses the placebo plus diet and physical activity arm of each trial as the common comparator. The CS and ITC report do not comment on the similarity of diet and physical activity prescriptions.

In both trials patients were advised to increase their physical activity to at least 150 minutes per week and adhere to a 500kcal deficit diet relative to their estimated individualised energy requirements. However, there were some differences between the trials, e.g. in the frequency and nature of the counselling sessions (individual sessions in STEP 1 every 4 weeks; individual or group sessions in SCALE 1839, frequency not reported).

Although the trials had different durations, as discussed above the company base their ITC analysis on outcomes measured approximately 1 year following randomisation. This was 52 weeks after randomisation for the STEP 1 trial (of which 36 weeks were on the full 2.4mg dose in the semaglutide 2.4 mg arm) and 56 weeks after randomisation for the SCALE 1839 trial (of which 52 weeks were on the full 3.0 mg dose in the liraglutide arm). However, the CS does not report outcomes for STEP 1 at 52 weeks after randomisation but instead reports them at 68 weeks after randomisation (the end of treatment). It is therefore not possible for the ERG to compare the outcomes in the placebo plus diet and physical activity arms of the trials at the same timepoints as used in the ITC.

The only comparison of the placebo plus diet and physical activity arms that the ERG can make based on the data provided by the company is for the FAS populations and the change from baseline to end of treatment, i.e. 68 weeks after randomisation in STEP 1 and 56 weeks after randomisation in SCALE 1839 (Table 24). NB the data reported in Table 24 are for the treatment policy estimand.

Table 24 Changes from baseline for outcomes at end of treatment in the placebo plus diet and physical activity arms of STEP 1 and SCALE 1839: FAS populations

Outcome, mean change from baseline	STEP 1 (68 weeks) placebo + DPA arm [a]	SCALE 1839 (56 weeks) Placebo + DPA arm [b]
Weight change	-2.61 kg	-2.8 kg

Proportional weight change		-2.41 %-points	-2.6 %-points
BMI change		-0.92 kg/m^2	-1.0 kg/m^2
Waist circumference change		-4.13 cm	-3.9 cm
HbA_{1c} change		-0.15 %-points	-0.06 %-points
Systolic blood pressure change		-1.06 mmHg	-1.50 mmHg
Ratio to baseline [c]	HDL cholesterol	1.01	0.7
	Total cholesterol	1.00	1.0

DPA: diet and physical activity
[a] Source: data as reported in section 3.2.5. above, using the treatment policy estimand
[b] Source: trial publication[19]
[c] analysis based on log scale and geometric means

CS section B.3.3.1.3 states that the placebo arms of the two trials were very similar in terms of baseline characteristics but did produce slightly different results for change from baseline in BMI and other risk factors. The effects of placebo plus diet and physical activity do appear broadly similar for both trials, with the changes in outcomes from baseline being generally consistent across the trials in their direction and magnitude (Table 24). The decrease in weight and BMI was marginally smaller in the STEP 1 placebo plus diet and physical activity arm; however, there is uncertainty in how applicable these FAS results are to the population subgroup and timepoints analysed in the ITC.

3.3.5 Risk of bias assessment for studies included in the ITC

The company used seven criteria to assess the risk of bias for the two studies, SCALE 1839 and STEP 1, included in the ITC (CS Appendix Table 9). The ERG independently assessed the studies using the same criteria as the company and our judgements are reported in Appendix 1. Overall, the ERG consider both trials to be of good methodological quality but the risk of attrition bias is unclear in both trials. The reasons for the risk of attrition bias being unclear to the ERG in the STEP 1 and SCALE 1839 trials are:

- The company provide data which show some systematic differences in baseline characteristics between patients with observations and those with missing data (clarification response A13, attachment E, Tables 26 to 53) for the liraglutide-eligible subgroup. Patients with missing data had a mean age that was 2.5 to 4.8 years lower (treatment policy estimand) or 3.3 to 4.0 years lower (trial product estimand) than those who provided observations for analysis. Also, a lower proportion of the patients with missing data had dyslipidaemia and hypertension than those who provided data for analysis. It is unclear whether these differences would be clinically important and whether, after imputation, they would favour one trial over the other.

- The company did not provide a similar comparison of baseline characteristics for patients with missing/non-missing data for the FAS population.

ERG conclusion: The company's inclusion of the STEP 1 and SCALE 1839 trials in the ITC is appropriate, although the ERG believe the STEP 3 trial should also have been included. The baseline characteristics of the STEP 1 and SCALE 1839 trials are broadly homogeneous, supporting the combining of these trials in an ITC. The risk of attrition bias is unclear in both trials, introducing uncertainty (of unknown magnitude and direction) around the efficacy outcome estimates from the ITC.

3.4 Critique of the indirect comparison and/or multiple treatment comparison

3.4.1 Overview of the ITC

Two relevant comparators were defined by the decision problem: standard management with diet and physical activity; and liraglutide. Whilst diet and physical activity formed the comparator arm of STEP 1 (along with placebo), an indirect treatment comparison (ITC) was required to compare semaglutide 2.4 mg to liraglutide 3.0 mg.

The population for the liraglutide ITC (BMI≥35, non-diabetic hyperglycaemia, and high risk of CVD) was aligned with TA664[4] (clarification responses A9-A12 and A18). As the company own both semaglutide 2.4 mg and liraglutide 3.0 mg, the analysis methodology was informed by the company's access to individual patient data (IPD) for both the STEP 1 and SCALE 1839 trials.

A series of unadjusted and adjusted analyses were conducted for the ITC, as explained in section 3.4 below). The following outcomes were included:

- Change from baseline – continuous outcomes:
    - Body weight (%)
    - Waist circumference
    - HbA1c
    - Systolic blood pressure
    - Fasting HDL cholesterol
    - Fasting total cholesterol
- Dichotomous outcomes:
    - Type 2 diabetes incidence
    - Proportion achieving normoglycaemic status

The type 2 diabetes incidence endpoint was not reported due to too few events (ITC report, section 3.2). BMI was not included; no explanation is provided in the CS. The economic model, however, does not use this outcome. The model uses % change in body weight, which is synonymous to % change in BMI.

As noted in section 3.2.4.3, two estimands were employed, which differ in how they address intercurrent events (rescue medication use and treatment discontinuation). The base case in the ITC uses the treatment policy estimand, that is, participants were included irrespective of whether they used rescue medications or discontinued treatment. This may be viewed as the more conservative approach as it would not adjust for a higher use of rescue medications in the comparator arm and discontinuations in the treatment arm, both of which might be expected to favour the comparator arm.

The trial product estimand was used as a scenario analysis in the ITC. The aim of this analysis is to reduce bias arising from differences in treatments and dropouts between trial arms by adjusting, through imputation methods, for use of rescue medications, treatment switching, or treatment discontinuation. The trial product estimand analysis produced similar outcomes to the treatment policy estimand, and where there were differences generally the treatment policy results are the more conservative (section 3.5).

The ERG requested band plots of the use of rescue medications, and discontinuations over time for both STEP 1 and SCALE 1839 (akin to Figure 3 in Aroda et al 2019[18]) to try to understand how the incidence of these intercurrent events differed by treatment arm and between trials and thus impacted the estimands (clarification question A16). The company did not provide these, arguing that they would not provide a complete picture of STEP 1 due to subjects being able to discontinue then resume treatment. It is unclear to the ERG how this could not be presented as a band plot. Nevertheless, the STEP 1 data provided by the company show a higher use of rescue medications in the diet & physical activity arm (N=13 [1%] vs 7 [1%], clarification response document E, Table 25) and a higher rate of treatment discontinuation in the placebo (diet and physical activity) arm (~18% vs 12% at week 52, CSR Figure 14.1.11). It is unlikely the higher rate of placebo discontinuations would impact the results (reasons for discontinuations appeared broadly similar for both trials – see flow charts in CS Appendix and SCALE 1839 trial paper[19]).

For the SCALE 1839 trial, the company noted there was “no notion of anti-obesity rescue medication” nor any distinction between treatment discontinuation and trial withdrawal (clarification response A16). It is unclear to the ERG whether this means rescue

medications were not recorded or not permitted. The data provided (clarification responses, attachment W, Figure 14.1.7) show a higher rate of (any-cause) discontinuation in the liraglutide 3.0 mg arm compared to placebo (approximately 10% vs 3%).

The ERG were concerned that the approaches to handling missing data differed between trials (clarification question A13). The company clarified that the missing data approaches were equivalent for the treatment policy estimand across both trials but differed due to using pre-planned analyses for each of STEP 1 and SCALE 1839 for the trial product estimand. In response to the ERG question, the company aligned the approach to the trial product estimand by applying the same mixed model for repeated measures as used for STEP 1 to SCALE 1839. This resulted in similar though slightly less favourable estimates in favour of semaglutide 2.4 mg for body weight (%), waist circumference, and glycaemic status (clarification responses document E, Tables 2-9).

Less imputation was required for STEP 1 using the treatment policy estimand (~10% semaglutide 2.4 mg, 16% diet and physical activity) compared to SCALE 1839 (~23% liraglutide 3.0 mg, ~31% diet and physical activity). By definition, the trial product estimand requires more imputation, but again this was less in STEP 1 (~22% semaglutide 2.4 mg, ~26% diet and physical activity) compared to SCALE 1839 (~32% liraglutide 3.0 mg, 33% diet and physical activity).

The differences in how or whether intercurrent events are recorded in the two trials raise questions about how they can be consistently handled in the missing data imputation used to calculate the estimands. The trial product approach

- requires more imputation, and therefore introduces more uncertainty than the treatment policy estimand
- relies on poorer recording of intercurrent events in SCALE 1839
    - There is no distinction between treatment discontinuations and other trial withdrawals (patients are therefore grouped together)
    - Use of rescue medications was not permitted or not recorded
- requires more imputation in SCALE 1839 than STEP 1

We conclude that the treatment policy estimand is likely to be a more conservative scenario for the efficacy of semaglutide 2.4 mg and is also less uncertain as less imputation is required. (As discussed in section 4.2.6.1, the economic model uses the trial product estimand from the STEP 1 trial and we believe this is appropriate for the purposes of the

economic model; note, however, that neither the treatment policy not trial product estimand results from the ITC have been used in the model.)

Whilst the Company's ITC base case compared outcomes at week 52 for STEP 1 versus week 56 for SCALE 1839, scenarios considered week 56 for both trials, and week 68 for STEP 1 versus week 56 for SCALE 1839. Results for these scenarios are relatively consistent across outcomes (section 3.3.3.2).

3.4.2 Data inputs for the ITC

The ERG agree that the patients' characteristics are relatively homogeneous between the semaglutide 2.4 mg and liraglutide populations (section 3.3.3.1 above). This applies both the FAS (STEP 1[5] and SCALE 1839[19] trial publications) and the liraglutide-eligible subgroup (CS Table 11). There were some minor differences in Asian and black/African American ethnicity (clarification response A11, attachment E, Table 1) but this was not identified as an effect modifier by the company and therefore not included in the adjusted analysis.

The company identified body weight, gender, baseline hbA_{1c}, and age as effect modifiers, and dyslipidaemia, hypertension, and CVD as potential effect modifiers (CS, section B.2.9.1.1). They note that neither race nor ethnicity were found to be effect modifiers for semaglutide or liraglutide (see section 3.3.3.2 above). The ERG's expert did not identify any missing effect modifiers.

3.4.3 Statistical methods for the ITC

The company conducted a series of adjusted and unadjusted analyses for the ITC. The adjusted analyses used established methods, linear regression (for continuous outcomes) and logistic regression (for dichotomous outcomes) to control for effect modifiers (body weight, gender and baseline hbA_{1c}) and potential effect modifiers (dyslipidaemia, hypertension, and CVD). Given the similarity in results and the similarity between the semaglutide 2.4 mg and liraglutide populations, the company preferred the unadjusted analysis as their base case (CS section B.2.9.1.2).

The ERG agree the unadjusted ITC is adequate to compare semaglutide 2.4 mg and liraglutide 3.0 mg. The semaglutide 2.4 mg and liraglutide 3.0 mg populations are homogeneous in terms of baseline characteristics and effect modifiers (CS Table 11) hence any adjusted ITC would not be expected to have a material impact on relative treatment effects.

The company provided the SAS code used for the ITC in clarification response A8. They declined to provide the IPD, and hence the ERG were unable to validate the adjusted ITC results. However, we were able to confirm the unadjusted ITC results using the Bucher method using the data reported in the ITC report, for all outcomes and both estimands.

Finally, neither the adjusted nor unadjusted ITC results inform the economic model. Instead, a separate ad hoc calculation was performed by the company to adjust for "slightly different results for change from baseline in BMI and other risk factors" (which were not specified) (CS section B.3.3.1.3). The company's calculation adjusts the efficacy of liraglutide 3.0 mg in the economic model to reflect this difference using observed efficacy in SCALE 1839. The mean changes from baseline in STEP 1 (trial product estimand) are used directly in the economic model (CS Table 21), whilst for liraglutide an odds ratio from SCALE 1839 was applied to the diet and physical activity arm of STEP 1 to give the adjusted estimates for liraglutide (CS Table 23). However, the details of this calculation are unclear to the ERG. As we note above, the differences in the changes from baseline for BMI and other outcomes between the diet and physical activity arms are relatively small (section 3.3.4), but the company do not provide a rationale for why the unadjusted ITC could not have been used in the economic model (i.e. avoiding the need for this ad hoc calculation). The company state in CS section 3.3.1.3 that the ITC was "not able to produce adjusted estimates for efficacy in responders (further details are provided in Appendix D)". However, there is no reference to this in CS Appendix D.

3.4.4 Summary of the ERG's critique of the ITC

- The ITC methodology followed by the company is appropriate given the available data.
- The methodology has been described and applied correctly.
- All effect modifiers have been included in the analysis.
- The adjusted ITC could not be validated as IPD were not provided.
- The unadjusted ITC results are preferred for the ITC since the STEP 1 and SCALE 1839 trial populations are homogeneous.
- A comprehensive range of scenario analyses were conducted by the company.
- The treatment policy estimand (company ITC base case) is likely to be the most conservative; the trial product estimand makes more use of data imputation which may introduce bias (or at least uncertainty) since missing data are inconsistently reported between trials. Use of the trial product estimand in the economic model is

appropriate, as it takes into account treatment stopping, which the treatment policy estimand does not (see section 4.2.6.1).

- It is unclear why the ITC results were not implemented in the economic model.
- The company's adjustment calculation in the economic model, used in lieu of relative effectiveness data from the ITC, is unclear to the ERG.

3.5 Results from the indirect comparison

The CS states (section B.2.9.2) that results of the unadjusted population analysis at the primary time point of interest using the trial product estimand, and the results of the scenario analyses (population adjustment 1, population adjustment 2, and unadjusted non-diabetic hyperglycaemia population), are provided in Appendix D4. However, Appendix D4 was not provided with the submission. The available ITC results presented below are from the ITC Report and CS Table 12. We report results below for the outcomes that inform the economic model (although, note, none of the ITC results were used in the model). We report the ITC results for other outcomes in Appendix 3.

3.5.1 Body weight

The unadjusted analyses for both the treatment policy estimand and trial product estimand indicate a statistically greater weight reduction with semaglutide 2.4 mg than with liraglutide 3.0 mg (Table 25). Adjusted analyses are only reported for the treatment policy estimand and these were also significantly in favour of semaglutide 2.4 mg. The treatment effect in unadjusted analyses was consistently larger for the trial product estimand than for the treatment policy estimand.

Table 25 ITC results: effect on % weight change from baseline

Analysis (STEP 1/SCALE 1839: week 52/56 unless stated)	**Relative treatment effect (95% CI), %-points semaglutide 2.4 mg vs liraglutide 3.0 mg**	
	Treatment policy estimand	**Trial product estimand**
Unadjusted [a]	-5.81 (-7.62 to -3.99), p < 0.0001 [a,b]	-6.62 (-8.28, -4.96), p<0.0001 [b]
Population adjustment 1	-5.87 (-7.69, -4.06), p<0.0001 [b]	Not reported
Population adjustment 2	-5.72 (-7.56, -3.89), p<0.0001 [b]	Not reported
Unadjusted, pre-diabetes	-5.78 (-7.06, -4.49), p<0.0001 [b]	Not reported
Week 56/56, unadjusted	-5.98 (-7.83, -4.14), p<0.0001 [b]	Not reported
Week 68/56, unadjusted	-6.51 (-8.51, -4.51), p<0.0001 [b]	-7.59 (-9.40, -5.79) [b]
Week 28/28, unadjusted	-2.92 (-4.22, -1.61), p<0.0001 [b]	-3.35 (-4.57, -2.13), p<0.0001 [b]
[a] From CS Table 12		

[b] From ITC Report Table 5

3.5.2 Glycaemic status

The CS states that semaglutide 2.4 mg was associated with a statistically significantly higher odds of achieving normo-glycaemic status compared to liraglutide 3.0 mg (CS section B.2.9.2). However, the odds ratio was not statistically significant for all the analyses conducted (Table 26). Notably, the primary unadjusted analysis (week 52 in STEP 1 compared against week 56 in SCALE 1839) was only marginally significant for the treatment policy estimand analysis, with the lower limit of the 95% confidence interval of the odds ratio being fractionally above 1.0). The odds ratio for the trial product estimand analysis was higher and statistically significant, but with a relatively wide 95% confidence interval. Odds ratios for the adjusted analyses were reported only for the treatment policy estimand analysis and were not statistically significant.

The CS states that the lack of a difference after adjusting for trial populations "was driven by a slightly lower baseline HbA1c in SCALE 1839 (5.8%) versus STEP 1 (5.9%); the closer a population is to being normo-glycaemic (i.e. HbA_{1c} < 5.7%), the lower the incremental glycaemic effect of adding a more potent GLP-1 receptor agonist" (CS section B.2.9.2).

Table 26 ITC results: effect on normoglycaemic status change from baseline

Analysis (STEP 1/SCALE 1839: week 52/56 unless stated)	**Odds ratio (95% CI), semaglutide 2.4 mg vs liraglutide 3.0 mg**	
	Treatment policy estimand	**Trial product estimand**
Unadjusted	1.79 (1.01, 3.16), p=0.0455 [a,b]	2.36 (1.26, 4.43), p=0.0073 [b]
Population adjustment 1	1.52 (0.82, 2.79), p=0.1804 [b]	Not reported
Population adjustment 2	1.56 (0.84, 2.92), p=0.1618 [b]	Not reported
Unadjusted, pre-diabetes	1.61 (1.07, 2.41), 0.0220 [b]	Not reported
Week 56/56, unadjusted	1.86 (1.05, 3.29), p=0.0327 [b]	Not reported
Week 68/56, unadjusted	1.32 (0.76, 2.30), p=0.3263 [b]	2.44 (1.30, 4.60), p=0.0055 [b]
Week 28/28, unadjusted	2.03 (1.13, 3.65), p=0.0178 [b]	1.86 (1.03, 3.38), p=0.0405 [b]

[a] From CS Table 12
[b] From ITC Report Table 11

3.5.3 Systolic blood pressure

There was no statistically significant effect of semaglutide 2.4 mg compared to liraglutide 3.0 mg on systolic blood pressure, apart from in an unadjusted analysis for the treatment policy estimand in a prediabetes subgroup (Table 27). The CS comments that although differences were not significant, the reduction in SBP was numerically greater with semaglutide 2.4 mg than with liraglutide 3.0 mg (CS section B.2.9.2). We note that the difference in all analyses was very small, in all cases less than 3.0 mmHg.

Table 27 ITC results: effect on systolic blood pressure change from baseline

Analysis (STEP 1/SCALE 1839: week 52/56 unless stated)	**Relative treatment effect (95% CI), mmHg, semaglutide 2.4 mg vs liraglutide 3.0 mg**	
	Treatment policy estimand	**Trial product estimand**
Unadjusted	-1.64 (-4.60, 1.32), p=0.2783 [a,b]	-1.36 (-4.04, 1.32), p=0.3197 [b]
Population adjustment 1	-1.92 (-4.87, 1.04), p=0.2032 [b]	Not reported
Population adjustment 2	-1.59 (-4.53, 1.34), p=0.2874 [b]	Not reported
Unadjusted, pre-diabetes	-2.82 (-4.89, -0.74), p=0.0078 [b]	Not reported
Week 56/56, unadjusted	-1.56 (-4.32, 1.20), p=0.2672 [b]	Not reported
Week 68/56, unadjusted	-1.32 (-4.25, 1.60), p=0.3751 [b]	-1.26 (-3.88, 1.37), p=0.3477 [b]
Week 28/28, unadjusted	-1.36 (-4.25, 1.54), p=0.3582 [b]	-1.55 (-4.33, 1.22), p=0.2730 [b]

[a] From CS Table 12
[b] From ITC Report Table 6

3.5.4 Fasting HDL and total cholesterol

The CS concludes that semaglutide 2.4 mg and liraglutide 3.0 mg resulted in similar changes from baseline in HDL and total cholesterol (CS section B.2.9.2). This is corroborated by results reported in the CS and ITC Report for HDL cholesterol (Table 28) and for total cholesterol (Table 29). We note that the change from baseline in HDL and total cholesterol was very small, with ratios to baseline being very close to 1.0 for semaglutide 2.4 mg[5] and change from baseline being ≤3.1 %-points for liraglutide 3.0 mg.[19]

Table 28 ITC results: effect on fasting HDL cholesterol change from baseline

Analysis (STEP 1/SCALE 1839: week 52/56 unless stated)	**Ratio to baseline (95% CI), semaglutide 2.4 mg vs liraglutide 3.0 mg**	
	Treatment policy estimand	**Trial product estimand**
Unadjusted	1.01 (0.98, 1.04), p=0.5696 [b]	1.01 (0.98, 1.04), p=0.5843 [b]

Population adjustment 1	1.01 (0.98, 1.04), p=0.4430 [b]	Not reported
Population adjustment 2	1.01 (0.98, 1.04), p=0.5028 [b]	Not reported
Unadjusted, pre-diabetes	1.00 (0.98, 1.02), p=0.9010 [b]	Not reported
Week 56/56, unadjusted	Not reported	Not reported
Week 68/56, unadjusted	1.03 (1.00, 1.07), p=0.0523 [b]	1.04 (1.00, 1.07), p=0.0437 [b]
Week 28/28, unadjusted	0.97 (0.94, 1.00), p=0.0261 [b]	0.96 (0.94, 0.99), p=0.0146 [b]
[a] p-value is from an updated version of the ITC Report and differs from that reported in CS Table 12 and the original version of the ITC Report (clarification response A17) [b] From ITC Report Table 7		

Table 29 ITC results: effect on fasting total cholesterol change from baseline

Analysis (STEP 1/SCALE 1839: week 52/56 unless stated)	**Ratio to baseline (95% CI), semaglutide 2.4 mg vs liraglutide 3.0 mg**	
	Treatment policy estimand	**Trial product estimand**
Unadjusted	0.97 (0.94, 1.00), p=0.0961 [a,b]	0.96 (0.93, 1.00), p=0.0278 [b]
Population adjustment 1	0.97 (0.94, 1.00), p=0.0955 [b]	Not reported
Population adjustment 2	0.97 (0.94, 1.00), p=0.0857 [b]	Not reported
Unadjusted, pre-diabetes	0.96 (0.94, 0.98), p=0.0004 [b]	Not reported
Week 56/56, unadjusted	Not reported	Not reported
Week 68/56, unadjusted	0.99 (0.95, 1.02), p=0.4096 [b]	0.98 (0.94, 1.01), p=0.1584 [b]
Week 28/28, unadjusted	0.97 (0.94, 1.00), p=0.0741 [b]	0.96 (0.93, 1.00), p=0.0261 [b]
[a] From CS Table 12 [b] From ITC Report Table 8		

3.6 Additional work on clinical effectiveness undertaken by the ERG

As we suggest the STEP 3 trial should have been included in the CS (see discussion in section 3.2.1), we have summarised results from the STEP 3 trial in Table 34 for outcomes that are used in the economic model.

Table 30 Summary of selected STEP 3 trial results

Estimand	Semaglutide 2.4 mg + IBT		Placebo + IBT		Difference (95% CI)
	Mean change	N	Mean change	N	
% body weight reduction					
Treatment policy	-16.0	407	-5.7	204	-10.3 (-12.0 to -8.6); p<0.001
Trial product	-17.6	407	-5.0	204	-12.7 (-14.3 to -11.0), p <0.001
Systolic blood pressure, mm Hg					
Treatment policy	-5.6	407	-1.6	204	-3.9 (-6.4 to -1.5); p = 0.001
Trial product	-6.21	407	-3.47	204	-2.74 (-5.12 to -0.36), p = 0.02
Total cholesterol					
Treatment policy	-3.8	407	2.1	204	-5.8 (-8.4 to -3.2); p < 0.001
Trial product	-4.5	407	2.1	204	-6.4 (-8.8 to -4.0), p < 0.001
HDL cholesterol					
Treatment policy	6.5	407	5.0	204	1.5 (-1.8 to 4.9), p = 0.39
Trial product	6.2	407	6.5	204	0.2 (-2.5 to 3.0), p = 0.860
Percentage of participants with prediabetes or non-diabetic hyperglycaemia at baseline who achieved normoglycaemia					
Treatment policy	NR	NR	NR	NR	NR
Trial product	NR	NR	NR	NR	NR

Source: Wadden et al. (2021)[3]

IBT: intensive behavioural therapy

3.7 Conclusions on the clinical effectiveness evidence

The company provided evidence in the CS that compares the clinical efficacy of semaglutide 2.4 mg in addition to standard weight management against two of the three comparators specified in the NICE scope:

- Standard management without semaglutide – referred to as 'diet and physical activity' in this report.
- Liraglutide (for the population for whom liraglutide is recommended in technology appraisal 664[4]), i.e. people with a BMI of ≥ 35 kg/m^2 with non-diabetic hyperglycaemia and a high risk of cardiovascular disease. In the included liraglutide trial (SCALE 1839) liraglutide was administered as an adjunct to standard management.

The company did not include the NICE scope specified comparator orlistat, and we believe that this is reasonable (see section 2.3).

The population specified in the NICE scope was adults living with overweight (BMI ≥ 27 kg/m^2 to < 30 kg/m^2) who had at least one comorbidity or living with obesity (BMI ≥ 30 kg/m^2). In their decision problem, the company have focused on a sub-population of the scope-specified population (see section 2.3). The company focus on people with a BMI of ≥ 30 kg/m^2 with at least one weight-related co-morbidity. However, the CS provides trial results for both this subgroup and the scope-specified population.

The company included one trial in their review, STEP 1, that directly compared semaglutide 2.4 mg plus standard management against a placebo arm that included standard management without semaglutide 2.4 mg. The STEP 1 trial participants were those living with overweight (BMI ≥ 27 kg/m^2) or obesity (BMI ≥ 30 kg/m^2) who had at least one weight-related co-morbidity. People with type 2 diabetes were not included in the trial. The company provided trial results for the full analysis set (n = 1,961) and two post-hoc subgroups: people with a BMI ≥ 30 plus at least one weight-related comorbidity (n = 1,470) (the 'target subgroup') and people with a BMI ≥ 35 with non-diabetic hyperglycaemia and high CVD risk (n = 421) (the liraglutide-eligible subgroup).

Regarding the representativeness of the STEP 1 trial participants' baseline characteristics, expert advice to the ERG was that in clinical practice, people with a higher BMI than those included in the trial are typically seen in tier 3 weight management services and thus people have more comorbidities. The trial may therefore not be fully representative of the people treated in practice in these respects. We considered the trial to have been generally well-

conducted, but the ERG are unclear about the risk of attrition bias which introduces some uncertainty (of unknown magnitude and direction) to the outcome estimates reported in the CS.

The trial found participants treated with semaglutide 2.4 mg showed a consistently higher percentage decrease in weight from baseline at 68 weeks than those treated with standard management in the FAS population and both subgroups. The proportion of patients with prediabetes or non-diabetic hyperglycaemia at baseline who achieved normoglycaemia at week 68 was higher for the semaglutide 2.4 mg arm than with standard management for the FAS and liraglutide-eligible groups. This outcome was not reported for the target subgroup. There were greater improvements in systolic blood pressure from baseline up to week 68 semaglutide 2.4 mg than with standard management in the FAS population and both subgroups. In the FAS population, changes in HDL and total cholesterol from baseline to week 68 favoured semaglutide 2.4mg. HDL and total cholesterol results were not reported for the target and liraglutide-eligible subgroups.

In terms of adverse events, gastrointestinal disorders were more common in the semaglutide 2.4 mg plus standard management arm than with standard management (74.2% versus 47.9%). There were three cases of acute pancreatitis in the semaglutide 2.4 mg plus standard management arm (0.2%), versus none with standard management alone.

The ERG have identified the following concerns and uncertainties about the decision problem and clinical effectiveness evidence included in the CS for the comparison of semaglutide 2.4 mg against diet and physical activity:

- We consider that the company's focus on the BMI ≥ 30 plus at least one weight-related comorbidity target subgroup in their decision problem is acceptable (see section 2.3 for a discussion about this). We suggest, however, that the NICE criteria for eligibility for bariatric surgery may more suitably define the target population (BMI ≥ 35 with at least one co-morbidity or ≥ 40 with or without comorbidities, unless new onset diabetes, in which case BMI ≥ 30, or lower for people of Asian family origin).[1] These criteria reflect the patient group that is typically treated within tier 3 services where we understand semaglutide 2.4 mg is most likely to be used.
- It is uncertain if the STEP 2 trial of semaglutide 2.4 mg in people with type 2 diabetes should have been included in the CS. If it is expected that the 2.4 mg dose might be used in practice in people with type 2 diabetes for weight loss and maintenance, then

data from STEP 2 trial will be relevant. There are currently no efficacy data for this population in the submission.

- The company post-hoc excluded the STEP 3 trial from their review, as it used IBT as part of the standard care arm (i.e. alongside a reduced calorie diet and increased physical activity). We believe the trial meets the NICE scope and it would have been appropriate to include it in the review. Omission of it means it is uncertain how effective semaglutide 2.4 mg would be when all relevant evidence has been considered.
- Two completed semaglutide 2.4 mg trials (STEP 5 and STEP 8) are relevant to the NICE scope and the company states clinical study reports for these trials are expected in Q3 and Q4 of this year. These studies' data could potentially have a bearing on conclusions about the clinical effectiveness and cost-effectiveness of semaglutide 2.4 mg.

A further issue we note is that there are uncertainties around how long people should be treated with semaglutide 2.4 mg, given that obesity is a long-term condition, and whether treatment could be repeated.

The company conducted an ITC, using individual patient data, to compare semaglutide 2.4 mg to liraglutide 3.0 mg. The company included the STEP 1 and SCALE 1839 trials. The ITC utilised data from the liraglutide-eligible subgroup of patients, i.e. those with BMI≥35 kg/m^2 with non-diabetic hyperglycaemia and high CVD risk. Like the STEP 1 trial, we considered the SCALE 1839 trial to have been well-conducted but at risk of attrition bias.

The results indicated statistically significant greater weight reduction with semaglutide 2.4 mg than with liraglutide 3.0 mg. The CS states that semaglutide 2.4 mg was associated with a statistically significantly higher odds of achieving normo-glycaemic status compared to liraglutide 3.0 mg. However, the odds ratio was not statistically significant for all the analyses conducted. There was no statistically significant effect of semaglutide 2.4 mg compared to liraglutide 3.0 mg on systolic blood pressure, apart from in an unadjusted analysis for the treatment policy estimand in a prediabetes subgroup. Semaglutide 2.4 mg and liraglutide 3.0 mg resulted in similar changes from baseline in HDL and total cholesterol.

The ITC methodology followed by the company is appropriate, but we have identified the following concerns and uncertainties:

- It is unclear why the ITC results were not implemented in the economic model.

- The company's adjustment calculation in the economic model, used in lieu of relative effectiveness data from the ITC, is unclear to the ERG.

4 COST EFFECTIVENESS

4.1 ERG critique on the company's review of cost-effectiveness evidence

The company conducted a systematic literature review (SLR) to identify all relevant economic evaluation studies, and resource use and cost studies for adults with obesity (CS section B.3.1 and CS Appendix G). The company updated searches that had previously been conducted for liraglutide for NICE technical appraisal (TA664)[4] and conducted new searches related to semaglutide.

The company performed their searches in relevant electronic databases and conferences (CS Appendix G Table 16 and section G.2.) The searches were conducted in April 2021. The ERG note that Health Technology Assessment (HTA) databases were not searched. The inclusion and exclusion criteria are presented in CS Appendix G Table 27. The original inclusion criteria for TA664[4] were for patients treated with liraglutide, orlistat or usual care (diet and physical activity) and the new searches for this appraisal included patients treated with semaglutide. Studies were only included if they were conducted in the UK.

From the 58 publications that met the inclusion criteria, seven were included in the company's review of cost-effectiveness / cost utility studies. None of the studies were for treatment with liraglutide or semaglutide. More details of the studies are reported in CS section 3.1 and CS Appendix G.7.

ERG conclusion

The ERG considers the company's review may have missed some potentially useful studies because they have not included studies outside the UK and they do not appear to have searched the grey literature, e.g. HTA reports. Studies conducted outside the UK may have been useful if they had reported the cost-effectiveness for semaglutide or liraglutide. Nevertheless, the ERG considers the most relevant published publication to be the NICE appraisal for liraglutide (TA664)[4] and the company based their cost-effectiveness model on the one developed for TA664.

4.2 Summary and critique of the company's submitted economic evaluation by the ERG

4.2.1 NICE reference case checklist

The NICE reference case checklist for the company's model is shown in Table 31. The ERG considers that the company model meets all the criteria of the NICE reference case.

Table 31 NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Yes
Perspective on costs	NHS and PSS	Yes
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes, although would be better as 50 year time horizon.
Synthesis of evidence on health effects	Based on systematic review	Yes
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Yes
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Yes
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Yes
Equity considerations	An additional QALY has the same weight regardless of	Yes

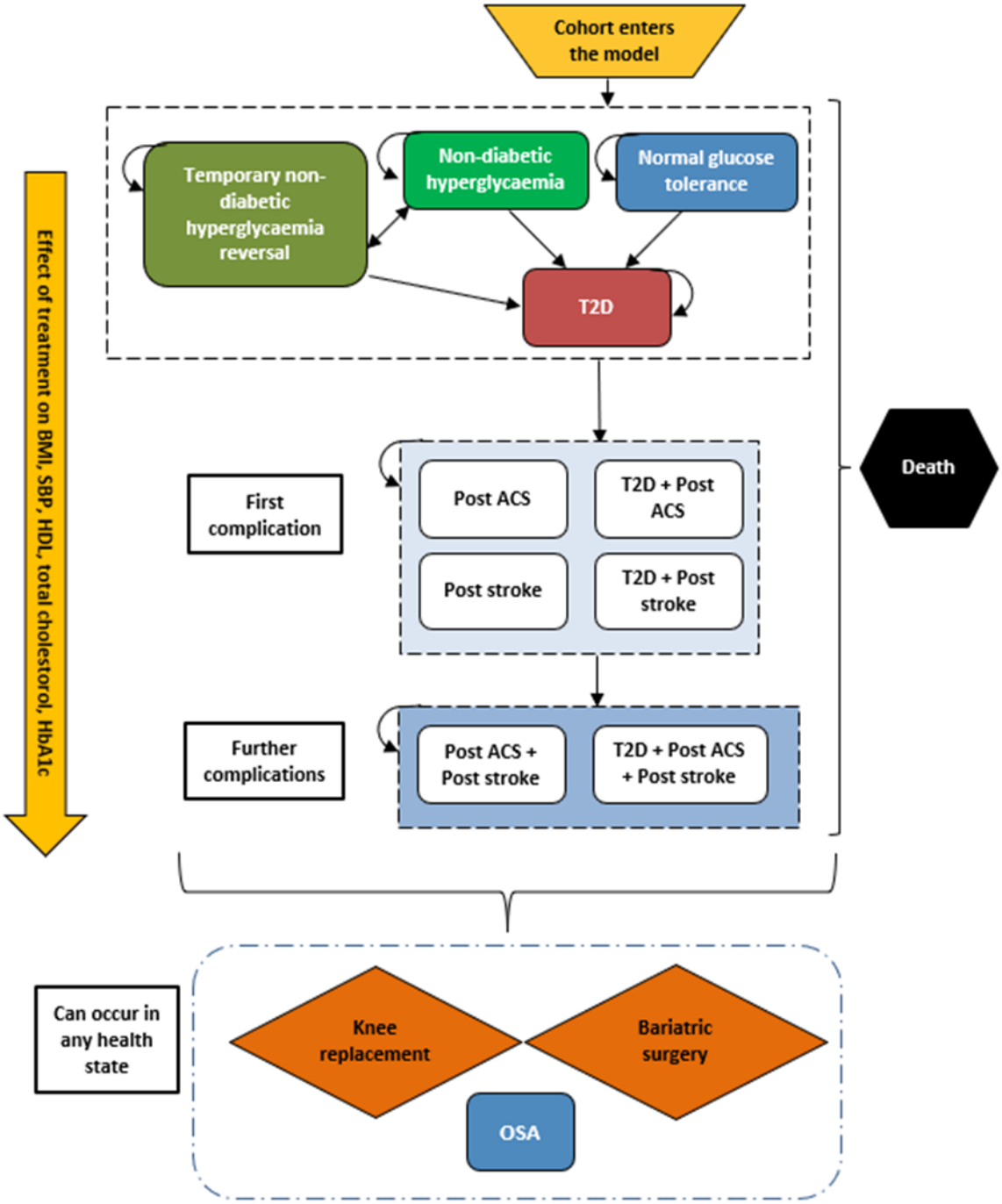
	the other characteristics of the individuals receiving the health benefit	
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Yes

4.2.2 Model structure

4.2.2.1 Overview of the model structure

The company developed a cohort state transition model using Microsoft Excel. The model structure is shown in Figure 1 (CS figure 11). The model was adapted from the model submitted to NICE as part of TA664 for liraglutide.[4] The model has a cycle length of three months for the first year, to allow for the incorporation of treatment discontinuation and then annual cycles thereafter. The model consists of 11 health states: Temporary non-diabetic hyperglycaemia reversal, Non-diabetic hyperglycaemia, Normal glucose tolerance (NGT), type 2 diabetes (T2D), Post acute coronary syndrome (ACS), Post stroke, T2D + post ACS, T2D + Post stroke, Post ACS + Post stroke, T2D + post ACS + post stroke. In addition to the health states, there are acute events that may occur from any health states for knee replacement, bariatric surgery and obstructive sleep apnoea (OSA).

Figure 1 Obesity model structure



Key: ACS, Acute coronary syndrome; BMI, body mass index; HDL, high density lipoprotein; OSA, obstructive sleep apnoea; SBP, systolic blood pressure; T2D, type 2 diabetes.
Source: reproduction of CS Figure 11.

Individuals enter the model in the NGT or non-diabetic hyperglycaemia health states, according to the observed prevalence in the STEP 1 clinical trial. In the base case analysis (BMI ≥ 30 kg/m^2 with one or more obesity related comorbidities) 46.6% of patients have NGT and 53.4% have non-diabetic hyperglycaemia. For the subgroup analysis all patients have non-diabetic hyperglycaemia.

During each model cycle, the cohort moves between health states or may remain in the same health state. The likelihood of transition between health states is given by the transition probabilities and these are calculated from risk functions using surrogate outcomes (BMI, SBP, total cholesterol, HDL cholesterol and HbA1c). The risk functions are described in more detail in section 4.2.6. Individuals in the non-diabetic hyperglycaemia health state may have a temporary reversal of their hyperglycaemia (in month 3) and move to the temporary reversal of non-diabetic hyperglycaemia health state. After cessation of treatment, individuals who remain in this state return to the non-diabetic hyperglycaemia health state.

Following an ACS event (non-fatal MI or non-fatal unstable angina event), individuals transition to a post ACS state. They transition to a post-stroke health state following a cerebrovascular event (non-fatal stroke or transient ischemia attack (TIA)). Those individuals in the post ACS health state remain in the same state if they experience further MI or unstable angina events in the next cycles. If they experience a cerebrovascular event they transition to the post ACS + post-stroke health state. Similarly, individuals residing in the post-stroke health state can experience a further stroke or TIA event in the next cycles but remain in the same post-stroke state or experience a cardiovascular event and transition to a post ACS + post-stroke health state. Patients in the cardiovascular health states described above are divided between those with and without diabetes. Transition to death can occur from any of the model health states either as a fatal event occurs or based on disease specific and general population mortality.

Patients with non-diabetic hyperglycaemia move to T2D + post-ACS or T2D + post stroke states following an ACS or stroke event respectively. The CS states that this assumption was previously used in TA664. The ERG notes that the NICE committee had reservations about this assumption and there was no good evidence to determine the proportion of people who develop type 2 diabetes after a cardiovascular event. The company includes a scenario analysis (CS Table 56) where patients do not develop T2D after a CVD event. Clinical advice to the ERG suggests that it is not possible to assume that patients will develop T2D after a CVD event. Therefore, in the ERG base case in section 6, we assume that patients do not develop T2D after a CVD event.

The benefits of treatment are introduced into the model through changes in the intermediate clinical outcomes (BMI, SBP, total cholesterol, HDL cholesterol) from the STEP 1 trial (described in more detail in section 4.2.6). Patients discontinue treatment after three months on the maintenance dose if they are a non-responder (defined as less than 5% weight loss

from baseline) or for other reasons such as adverse events. There is a maximum treatment duration of two years. Patients who discontinue treatment assume the same treatment efficacy as the diet and physical activity arm. Treatment effect is assumed to wane in a linear fashion over three years after discontinuation of treatment. Patients' clinical outcomes (BMI, SBP, total cholesterol and HDL cholesterol) gradually revert to be equal to those in the diet and physical activity arm and from there follow on the same path (CS Figure 12 and 13).

The model assumes a natural BMI increase after the treatment period of 0.1447 kg/m^2 for males and 0.1747 kg/m^2 for females per year for the diet and physical activity arm, as estimated by Ara et al.[24] After individuals reach 68 years, no further weight increase is assumed. Projected BMI change for semaglutide 2.4 mg and diet and physical activity over time is shown in CS Figure 12.

There are more recent studies estimating natural weight gain, such as Iyen et al. [25]. Iyen et al estimated BMI trajectories for a cohort of 264,230 individuals in the UK followed for 10 years. There was a mean increase of 1.06 kg/m^2 over 10 years. Zaninotto et al[26] explored BMI trajectories in the English Longitudinal Study of Ageing. They reported that after age 66 years, there was a steep decrease in individual's BMI, in contrast to the company's assumption of no change in weight after age 68 years. We conduct scenario analyses using these sources (section 6).

OSA is accounted for by calculating costs and quality of life decrements for the estimated prevalence of OSA each cycle. Osteoarthritis is not included in the model except for related to knee replacement. T2D microvascular complications are not included as distinct health states. For a proportion of patients with T2D, higher costs apply reflective of microvascular conditions.

Assumptions

An abridged summary of the main model assumptions is presented in Table 32. Assumptions related to risk functions are discussed in section 4.2.6.

Table 32 Summary of assumptions applied in the economic model

Analysis setting	Assumption / Setting	Company justification	ERG comments
Assumptions that differ from TA664[4]			
Comorbidities included	ACS T2D Stroke Sleep apnoea Osteoarthritis	Conservative assumption of limiting to the most economically significant comorbidities to reduce the number of health states and complexity. TA664 also included cancer health states. These were removed to reduce complexity and incorporate feedback provided by the ERG during TA664.	We agree it is reasonable to focus on those health states that are mostly impacted by reducing obesity and thus have the most impact on model results.
Mortality	Disease specific and BMI adjusted mortality (CPRD study)	Mortality was also adjusted by BMI in order to avoid underestimating the mortality and costs. In TA664 no adjustment for BMI was applied	We agree
Application of acute and health state disutilities	Acute event and health state disutilities are assumed to be additive.	Assumption, given existing evidence Gough et al. 2009 and TA664. Some disutilities have been updated since TA664 due to more appropriate sources identified in the literature searches.	Although TSD 12 recommends that disutilities should be multiplicative, we consider there is evidence that using additive disutilities are appropriate, see section 4.2.7.3.
Application of acute and health state costs	Acute event costs and health state costs are assumed to be additive.	Additive health state costs is in line with Ara et al. 2012[24] and TA664.[4] Cost sources were updated as a new targeted literature review were used, resulting in different cost inputs compared to TA664.	We agree.
Assumptions consistent with TA664			
Catch up rate for BMI and surrogate outcomes	Pharmacotherapy returns to value of natural progression in diet and physical activity at a constant rate of 33% per year	The application of a constant rate of 33% per year following treatment cessation is in line with Ara et al. 2012 and TA664 which assumed BMI	We consider the rate of weight gain after treatment cessation is uncertain as

Analysis setting	Assumption / Setting	Company justification	ERG comments
		returned to baseline value at 3 years after treatment cessation in a linear fashion.	there are no available follow-up in the STEP 1 or pharmacological weight loss clinical trials.
Natural weight increase after treatment discontinuation	Weight increase following Ara 2012 in the CPRD dataset, until the cohort reaches 68 years old in all treatment arms	Natural weight increase is a common assumption in obesity models supported by a model developed by NICE. Support is also found in the study by Heitmann and Garby, 1999 and the analysis on the UK CPRD.	There are more recent studies estimating natural weight gain, such as Iyen et al.[25] and Zaninotto.[26] We conduct scenario analyses using these sources (section 6).
Progression of SBP, total cholesterol and HDL cholesterol post-treatment and post waning of treatment effect periods	Post-treatment and waning of treatment effect, SBP, total cholesterol, and HDL cholesterol were assumed constant for the remainder of the time horizon.	For reasons of simplicity, the model only accounted for evolution based on the treatment effect. The cohort returns to baseline value, corresponding to the average in the cohort, which is then maintained over the entire time horizon of the model when treatment is discontinued. However, as the cohort is assumed to remain treated with antihypertensive medications, and accrues the cost of this, it is plausible to assume the averages would remain stable.	We agree.
Temporary reversal of non-diabetic hyperglycaemia to a NGT state, maintenance of the glucose status effect over time and risk of T2D in non-diabetic	All patients in the non-diabetic hyperglycaemia state were assigned a higher risk of developing T2D (vs NGT patients) by modification of the glycaemic status parameter in the corresponding T2D risk equations. In line with changes in glycaemic status	According to published risk equations, patients with non-diabetic hyperglycaemia have a higher risk of developing T2D than those with normal glucose tolerance. Changes in glycaemic status observed in STEP 1 and SCALE 1839 were applied in the model starting from Cycle 2.	We agree.

Analysis setting	Assumption / Setting	Company justification	ERG comments
hyperglycaemia vs NGT	observed in the STEP 1 and SCALE 1839 trials, a proportion of patients in semaglutide 2.4 mg, liraglutide 3.0 mg and diet and physical activity arms temporarily reverted to a normal glycaemic status whereby a lower risk of T2D was applied. All patients reverting to NGT were assumed to return to a non-diabetic hyperglycaemia status at the end of the treatment effect waning period at a constant rate of 33% per year, assuming glycaemic status be correlated with weight loss.	Non-diabetic hyperglycaemia reversal was assumed to be a consequence of the initial weight loss and thus applied in the model to occur at the same time. Consequently, the loss of temporary normo-glycaemia was also assumed to occur at the same time with the complete loss of the initial weight loss benefit.	
Stopping rule	Semaglutide 2.4 mg (if >5% weight loss not achieved at 28 weeks) Liraglutide 3.0 mg (if >5% weight loss not achieved at 16 weeks)	In line with anticipated semaglutide 2.4 mg marketing approval. In line with regulatory approval for liraglutide 3.0 mg.	We agree. We note that a stopping rule was not included in the STEP 1 clinical trial.
Treatment duration	2 years for semaglutide 2.4mg, liraglutide 3.0mg and diet and physical activity	This reflects clinical practice as weight management in SWMS is provided for two years. It is worth noting that after two years patients in the semaglutide 2.4 mg arm and the comparator arm transition to diet and physical activity alone because diet and physical activity is considered to be an integral part of lifelong weight management.	We agree.
Treatment discontinuation / retreatment	Patients who discontinued semaglutide 2.4 mg or liraglutide 3.0 mg treatment were assumed to remain on a diet and physical activity	Diet and physical activity was considered an integral part of the treatment of all individuals with obesity, regardless of any pharmacological or surgical intervention co-	We agree.

Analysis setting	Assumption / Setting	Company justification	ERG comments
	program for the rest of the analysis time horizon. It was assumed that there would not be any repeated course of treatment with pharmacotherapy	administered. No published clinical data was available to provide evidence with regards to a 'stop and re-start' type of weight management.	
Key: ACS, acute coronary syndrome; BMI, body mass index; CPRD, Clinical Practice Research Datalink; CV, cardiovascular; HDL, high-density lipoprotein; HRQoL, health-related quality of life; NGT, normal glucose tolerance; SBP, systolic blood pressure; T2D, type 2 diabetes; TLR, targeted literature review.			

Source: CS Table 51

ERG conclusion

The ERG considers that the model structure is appropriate and reasonable. It is based on the model submitted to NICE for the appraisal of liraglutide 3.0 mg for managing overweight and obesity (TA664) which was considered by the NICE committee to be suitable for decision making. The previous model also included cancer health states. The CS comments that this state was removed to reduce complexity and following comments of the ERG during TA664. There remains uncertainty around the rate of weight gain following treatment cessation and the assumption that patients with non-diabetic hyperglycaemia would develop type 2 diabetes after a cardiovascular event.

4.2.3 Population

The target population for the economic evaluation comprised of adult patients with a BMI of ≥ 30 kg/m^2 with at least one weight-related comorbidity (base case) and ≥ 35 kg/m^2 with non-diabetic hyperglycaemia and high risk for CVD. We refer to this subgroup in this report as the liraglutide-eligible subgroup.

The subgroup population is so defined in order to align with the recommended target population for liraglutide 3.0 mg (TA664).[4] The characteristics of the starting cohort for both populations are shown in Table 33 (CS Table 16) and were sourced from a post-hoc analysis of these subgroups in the STEP 1 clinical trial.

Table 33 Baseline characteristics for populations of interest

	Mean	
Patient characteristics	**BMI ≥ 30 kg/m^2 with one or more obesity related comorbidities**	**BMI ≥ 35 kg/m^2, with non-diabetic hyperglycaemia and high risk for CVD**
Used in comparison vs	Diet and physical activity	Liraglutide 3.0 mg
Age (years)	[redacted]	[redacted]
BMI (kg/m^2)	[redacted]	[redacted]
Height (m)	[redacted]	[redacted]
SBP (mmHg)	[redacted]	[redacted]
Total cholesterol (mg/dl)	[redacted]	[redacted]
HDL cholesterol (mg/dl)	[redacted]	[redacted]
HbA1c after T2D development (%)*	7.5	7.5
T2D duration (years)*	3.0	3.0
Triglycerides (mg/dl)	[redacted]	[redacted]
Proportion Triglyceride level >150 mg/dl (%)	[redacted]	[redacted]
Proportion current smokers (%)	[redacted]	[redacted]
Proportion females (%)	[redacted]	[redacted]
Proportion on lipid-lowering drug (%)	[redacted]	[redacted]
Proportion on antihypertensive medication (%)	[redacted]	[redacted]
Key: BMI, body mass index; HDL, high-density lipoproteins; KOL, key opinion leader; SBP, systolic blood pressure; T2D, type 2 diabetes. **Notes:** *Based on KOL opinion, applied after onset of diabetes. **Source:** STEP 1 trial [27]		

Source: Reproduction of CS Table 16

ERG conclusion

The population chosen by the company differs from that in the final scope issued by NICE, which is 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 at least one weight-related comorbidity. The CS states that a different population has been chosen as adult patients with a BMI of ≥ 30 kg/m^2 with at least one weight-related comorbidity are likely to benefit most from pharmacological treatment with SWMS in NHS clinical practice.

The ERG notes that tier 3 SWMS currently only see people with a BMI ≥35 kg/m^2 (obese); few people with a BMI of 30 to 35 kg/m^2 are treated within these services (see discussion in section 2.2.3). The ERG's clinical expert commented that that

people with higher BMIs than those included in the STEP 1 trial are typically seen in practice and thus people have more comorbidities (see section 3.2.1.2).

We also note the cost-effectiveness of semaglutide 2.4 mg is sensitive to the starting BMI of individuals (CS Figure 20). We have therefore presented results for different starting BMI cohorts in section 6**Error! Reference source not found.**.

4.2.4 Interventions and comparators

Semaglutide 2.4 mg is compared to diet and physical activity (for the population with BMI ≥ 30 kg/m^2 with at least one weight-related comorbidity) and liraglutide 3.0 mg (for the subgroup population with BMI ≥ 35 kg/m^2 non-diabetic hyperglycaemia and high risk for CVD).

The NICE scope also includes orlistat as a comparator which was not been included in the CS. The CS states that orlistat is not a relevant comparator for semaglutide 2.4 mg as it is no longer widely used in clinical practice. Further, in the NICE appraisal TA664,[4] orlistat was not considered as an alternative to liraglutide 3.0 mg by the NICE committee and was therefore not included as a comparator. As discussed in section 2.3, we agree it is reasonable to exclude orlistat as a comparator.

4.2.5 Perspective, time horizon and discounting

Costs are estimated from the NHS and Personal Social Services (PSS) perspective. Costs and QALYs are discounted at 3.5% in the base case (CS Tables 19). The model outcomes and costs are estimated over a lifetime horizon (40 years). Alternative time horizons of 20, and 30 years are considered in scenario analyses, but this assumption does not have a significant impact on the model results (CS Tables 56). We note that the most recent NICE appraisal for liraglutide 3.0 mg (TA664)[4] also applied a 40-year time horizon. The CS comments that there is a difference of less than 0.1% of patients alive between treatment arms after 40 years and therefore any subsequent differences beyond the modelled time horizon are expected to be minimal. The ERG considers that the lifetime horizon would be better for 50 years (until mean age of patients is 99 years), however we do not expect the results to change significantly with a longer time horizon.

ERG conclusion on perspective, time horizon and discounting

The company adopted the recommended perspective and discounting rates and an appropriate time horizon, which are all in line with NICE guidelines and previous NICE appraisals.

4.2.6 Treatment effectiveness and extrapolation

Transition probabilities between health states used in the model for T2D and cardiovascular events are based upon risk functions. The risk functions use intermediate endpoints, i.e. BMI, SBP, HDL and total cholesterol and HB1A1c to calculate transition probabilities. The treatment effect of the intervention is incorporated through a reduction in these intermediate clinical outcomes in the STEP 1 trial.

4.2.6.1 Treatment effect

The treatment effect is applied by a reduction in the clinical outcomes to the mean baseline values in each treatment arm. The outcomes used in the model are BMI, SBP, total cholesterol and HDL cholesterol and change in glycaemic status. The treatment effects for semaglutide 2.4mg and diet and physical activity for both subgroups were sourced from the FAS population of the STEP 1 clinical trial. The company state that this is reasonable as the population was defined a priori in the STEP 1 trial and therefore is a statistically robust measure of the treatment effect. The company conducted scenarios using the treatment effect specific to each of the subgroups, rather than using the FAS treatment effect (CS Table 56 and CS Table 60) and the results were similar to using the FAS treatment effect.

The trial product estimand is used in the model to reflect the treatment effect of those patient who remain on treatment. Note this differs from the treatment effect seen in the trial, reported in CS section 2.6. The model uses a stopping rule for non-responders which was not included in the STEP 1 trial. The trial product estimand used in the model is an adjusted treatment effect to incorporate this stopping rule. The trial product estimands for BMI (% weight change), SBP, total cholesterol and HDL cholesterol are shown in Table 34 (CS Table 21) for semaglutide 2.4 mg, diet and physical activity and liraglutide 3.0 mg (CS Table 23). The ERG considers the use of the trial product estimand to incorporate the effect of treatment discontinuation to be a reasonable and appropriate approach. The trial product estimand appears to be consistent with the treatment effect reported in the trial (see section 5.3). We note that the company STEP 1 trial did not include a stopping rule and there is some uncertainty about whether this will be included in the marketing authorisation for semaglutide 2.4mg (CS B3.2.3.1).

The treatment effect for liraglutide 3.0 mg was taken from the SCALE 1839 trial. In the absence of head-to-head data, the efficacy of liraglutide was adjusted based on the results of an indirect treatment comparison (discussed in section 3.4). The CS states that the adjustment was made to increase the estimated efficacy estimates of liraglutide by size of the difference between the efficacy estimates in the placebo arms (all patients) of STEP 1 and SCALE 1839.

Table 34 Change in physiological parameter values – STEP 1 and SCALE 1839 early responders

Parameter and timepoint in model	Semaglutide 2.4 mg: full analysis set N = 1306 Early responders n = [redacted]	Diet & physical activity: full analysis set N = 655	Liraglutide 3.0 mg: early responders N = 1456 (Week 28) Early responders n = [redacted] (Week 56)
	Mean change from baseline	Mean change from baseline	Mean change from baseline
Weight change (% change)			
Month 4	-12.04%	-2.69%	-10.00%
Month 7	-12.04%	-2.69%	-10.00%
Month 10	-13.22%	-2.44%	-10.00%
Year 1	-18.47%	-2.44%	-10.42%
Year 2	-18.47%	-2.44%	-10.24%
SBP change (mmHg)			
Month 4	-5.93	-0.56	-4.46
Month 7	-5.93	-0.56	-4.46
Month 10	-6.48	-1.14	-4.46
Year 1	-7.63	-1.14	-5.19
Year 2	-7.63	-1.14	-5.96
Total cholesterol change (mg/dl)			
Month 4	-15.27	1.39	-5.58
Month 7	-15.27	1.39	-5.58
Month 10	-15.92	0.18	-5.58
Year 1	-9.20	0.18	-2.25
Year 2	-9.20	0.18	-1.11
HDL cholesterol change (mg/dl)			
Month 4	-4.63	-0.96	-3.34
Month 7	-4.63	-0.96	-3.34
Month 10	-4.76	1.07	-3.34
Year 1	2.97	1.07	2.07
Year 2	2.97	1.07	2.17

Key: HDL, high density lipoprotein; SBP, systolic blood pressure.
Note: Early responders are defined as patients that achieve more than 5% weight loss at 28 weeks

Source: CS Table 21 and Table 23

The company assumes that the effect of treatment on BMI, surrogate outcomes (SBP, HDL and total cholesterol) and glycaemic status reduces linearly over three years after treatment cessation. At the end of three years, clinical parameters have returned to the same level as patients who received diet and physical activity alone. After the first model cycle (3 months), a proportion of patients with non-diabetic hyperglycaemia have glycaemic status reversal (90.4% for those treated with semaglutide 2.4 mg, 45.8% for diet and physical activity and 83.6% for liraglutide 3.0 mg). These patients will revert back to a non-diabetic hyperglycaemic status at the end of the treatment effect waning period at a constant rate of 33% per year. The proportions with glycaemic reversal for semaglutide 2.4 mg and diet and physical activity has been adjusted with an odds ratio between liraglutide and placebo (all patients) in the SCALE 1839.

4.2.6.2 Estimation of transition probabilities (risk equations)

The risk equations used to estimate transition probabilities between health states and for acute events are shown in Table 35 (CS Table 30). The risk equations use the cohort's mean clinical parameters combined with coefficients to estimate the risk of T2D and cardiovascular events. Where variables in the risk equations are not included as cohort characteristics or surrogate outcomes in the model, the average values of the derivation cohort of the risk equations were used and maintained constant over the time horizon of the analysis. The QRisk3,[28] QDiabetes[29] and UKPDS82[30] were large UK-based studies. The risk equations are described in more details in CS Appendix L.

Table 35 Risk equations used for obesity-related complications

Complication	Risk equation(s) available in model	Company justification for base case selection
Onset of T2D	QDiabetes-2018 Model C [29]	QDiabetes allows prediction of 10-year risk and includes BMI and HbA1c as predictive variables. This is in line with assumptions from TA664. [4]
	Framingham Offspring (scenario) [31]	
First CVD event	Qrisk3[28]	QRisk3 was estimated from a UK cohort and as such is being used in UK. This is in line with assumptions from TA664.[4]
	Framingham Heart Study (scenario) [32]	
Recurrent event	Framingham Recurring Coronary Heart Disease[32]	The only risk equation identified for recurrent CVD events in non-diabetic

		patients. This is in line with assumptions from TA664.[4]
First CVD event in T2D	UKPDS82[30]	The UKPDS 82 risk model (outcome model 2) is a large UK study and able to predict both first and recurrent CVD events after the onset of T2D. This is in line with assumptions from TA664[4]
	QRisk3[28]	
Incidence of recurrent CVD event in T2D	UKPDS82[30]	
	Framingham Recurring Coronary Heart Disease (scenario) [32]	
Onset of OSA	Sleep Heart Study[33]	This study was preferred to other available studies because it was the largest in sample size (n=5,615), it provided sufficient data to calculate a prevalence rate per unit BMI, and it investigated the prevalence of moderate-to-severe OSA (AHI ≥15), given that in the present health-economic analysis, OSA was assigned a hospital cost for continuous positive airway pressure treatment.
Knee replacement	Wendelboe et al. 2003[34]	The study provided granular data on the association between BMI and incidence of knee surgeries by 2.5 BMI-unit steps for observed BMI levels between 17.50 and 42.49 kg/m^2.

Source: reproduction of CS Table 30

Based on the QDiabetes equation the risk of developing T2D is higher in the cohort with non-diabetic hyperglycaemia than the NGT cohort. The risk factors for T2D included in the QDiabetes algorithm are shown in CS Appendix Table 58. HBA1c levels (blood sugar levels) were set constant in the model at 37 mmol/mol in NGT and 47 mmol/mol in non-diabetic hyperglycaemia. The ERG notes that HBA1c levels were recorded in the STEP 1 study (CS section 2.6.6).

The QRisk3 equation is used to derive the absolute risk of CVD in primary prevention in the base case. The risk factors for CVD in primary prevention used in the QRisk3 prediction model are shown in CS Appendix Table 61. The model stratifies the risk of CVD into the risk of angina, MI, stroke and TIA according to the proportions shown in CS Table 31.

The UKPDS 82 risk model is used to predict both first and recurrent CVD events after the onset of T2D. The risk factors and coefficients from UKPDS82 for CVD in primary prevention in the T2D cohort are shown in CS Appendix 64 and for 2nd MI and 2nd stroke are shown in CS Appendix Table 67. Generally, the ERG prefers the QRisk3 risk model to UKPDS 82 as this study is more recent so it is our view that the QRisk3 risk model should be used for the

prediction of CVD in individuals with T2D. This is also consistent with the risk model used for 1st CVD events for individuals without T2D.

Sleep apnoea is included as a comorbidity in the model with the prevalence of OSA dependent on BMI according to the Sleep Heart Health Study. Sleep apnoea prevalence by BMI level is shown in CS Appendix Table 48.

The risk of knee replacement is stratified by BMI, gender and age (<65; >65 years). The annual incidence was sourced from the study of Wendelboe et al.[34]. The risk functions for knee osteoarthritis are shown in CS Table 32.

Bariatric surgery is an option for people with severe obesity in whom non-surgical interventions have been tried but the individuals did not achieve the required weight loss. The criteria for bariatric surgery in the England is BMI ≥ 40 kg/m^2 or BMI ≥ 35 kg/m^2 with at least one comorbidity such as T2D or CVD. The company assumes the proportion of eligible patients undergoing bariatric surgery is 1.15% per year.[1] The model treats bariatric surgery as an annual event for patients who fulfil the above criteria. There are three types of bariatric surgery currently used available: gastric bypass, laparoscopic banding and sleeve gastrectomy. The prevalence of the types of bariatric surgery and their efficacy are shown in CS Table 27.

4.2.6.3 Mortality

Patients may die from any health state and mortality is split into short (associated with acute events) and long-term mortality. Mortality from CVD events, knee replacement and bariatric surgery are shown in CS Table 34. General population mortality from UK lifetables[35] was adjusted by excluding mortality of obesity related complications. Mortality was adjusted by a hazard ratio for BMI, based on Bhaskaran et al.[36] For patients in the Post ACS and Post stroke health states, a relative risk was applied to the mortality rate (CS Table 33).

The model does not include a hazard ratio for Type 2 diabetes mortality. A study by Mulnier et al[37] followed a cohort of patients with and without diabetes from the General Practice Research Database. They found higher mortality in individuals with diabetes than without. The HR for all-cause mortality in Type 2 diabetes compared with no diabetes was 1.93. However, we consider this HR for all-cause mortality may already be included within the hazard ratio for BMI and post-MI and post stroke and so we consider it is appropriate not to include a separate HR for Type 2 diabetes.

4.2.6.4 Treatment discontinuation

Treatment discontinuation occurs in the model due to:

i) Per cycle discontinuation due to any reason, such as adverse events,
ii) Non-responder early discontinuation or stopping rule,
iii) Maximum treatment duration of two years.

The probability of discontinuation per cycle was taken from the Kaplan-Meier curve of time to discontinuation.

The non-responder discontinuation / stopping rule applies to individuals who do not lose 5% of their initial body weight after 12 weeks of the maintenance dose. Thus, the stopping rule occurs after 28 weeks for semaglutide (16 weeks titration period, 12 weeks maintenance dose) and 16 weeks for liraglutide (4 weeks titration period, 12 weeks maintenance dose). The CS states that the company is expecting the marketing authorisation for semaglutide 2.4 mg to include a stopping rule that treatment would be discontinued for non-responders after 12 weeks on the maintenance dose, as is the case for liraglutide 3.0 mg.

4.2.6.5 Adverse events

Treatment related adverse events are included in the model for non-severe hypoglycaemia and severe gastrointestinal events. The incidence of the adverse events during the treatment period are shown in CS Table 25. AE rates were sourced from the STEP 1 trial for semaglutide 2.4 mg and the SCALE 1839 trial for liraglutide 3.0 mg.

ERG conclusion on treatment effectiveness and extrapolation

The company uses a trial product estimand for those patients who remain on treatment. The ERG considers the use of the trial product estimand to incorporate the effect of treatment discontinuation to be reasonable and appropriate. The company has provided validation to show that the trial product estimand appears to be consistent with the treatment effect reported in the trial (see section 5.3). We note that the company STEP 1 trial did not include a stopping rule and there is some uncertainty about whether this will be included in the marketing authorisation for semaglutide 2.4mg (CS B3.2.3.1).

The company used risk equations to estimate the long-term risk of morbidity. We consider that the use of these risk equations is appropriate to model diabetes and

cardiovascular outcomes based on surrogate outcomes. The same risk equations were used as in TA664. For that appraisal, the NICE committee accepted that the risk equations selected in the company's and ERG's base case were both suitable for decision making.

4.2.7 Health-related quality of life

4.2.7.1 Systematic literature review for utilities

The company conducted a systematic literature search to identify HRQoL studies for adults with obesity by updating searches that had previously been conducted for liraglutide for NICE technical appraisal (TA664)[4] and conducting new searches related to semaglutide. The search strategy is described in CS Appendix H. The searches were conducted on 14 April 2021. The company searched relevant database (see CS Appendix Table 32 and 33).

The eligibility criteria for the HRQoL studies included adults with BMI ≥35 kg/m^2 with treatment with liraglutide, orlistat, usual care (diet and physical activity) in the original review for TA664 and semaglutide for the current appraisal. (CS Appendix Table 44 and 45). The searches identified 26 unique HRQoL studies after abstract and full-text screening. The company comments that the studies identified were not relevant for the current economic analysis and were therefore not used.

4.2.7.2 Study-based health-related quality of life

The SF-36 and IWQOL-lite-CT HRQoL measures were collected from patients in the STEP 1 trial. These are reported in more detail in CS section B 2.6.9. The company comments that these measures are not consistent with the NICE reference case and do not yield utilities that could be used in the economic model.

4.2.7.3 Health-related quality of life data used in the company's cost-effectiveness analysis

The approach the company used for health-related quality of life was to use a baseline utility for individuals with no complications (in the normal glucose tolerance, non-diabetic hyperglycaemia and temporary non-diabetic hyperglycaemic reversal health states), based on age and BMI. Individuals in other health states or who suffer an acute event (such as stroke, TIA or knee replacement) are assigned a utility decrement associated with that health state or event.

4.2.7.3.1 Baseline utility

The company uses the study by Søltoft et al[38] for the baseline utility values. This study analysed the EQ-5D responses of 14,416 adults in the 2003 Health Survey for England. The company chose this study because it adjusted the utilities so that they are free of any additional obesity related comorbidities and the utility values are reported with coefficients for age and BMI. The ERG consider this to be an appropriate approach.

The utility curves related to BMI from Søltoft et al[38] are shown in CS Appendix Figure 15. The company notes that utility values appear to decline linearly after a BMI level of 25 kg/m^2. They therefore fit a linear function to the curve after this point. The coefficients used for utility based on BMI are shown in Table 36 (CS Table 35).

Table 36 Coefficient used to estimate baseline utility values based upon BMI

Parameter	(BMI 15-35 kg/m^2)		(BMI 36 kg/m^2 and beyond)	
	Males	**Females**	**Males**	**Females**
BMI3	0.000033	0.000017	.	.
BMI2	-0.003200	-0.001800	.	.
BMI	0.099000	0.057200	-0.105431	-0.147297
Constant	-0.020554	0.401769	1.323834	1.462846
Key: BMI, body mass index.				

Source: reproduction of CS Table 35

In a similar way utility values were adjusted for age based on the coefficients reported in Søltoft et al[38]. The coefficients from the Søltoft et al study are shown in CS Table 36.

4.2.7.3.2 Decrements in utility associated with acute events

Non-fatal acute events considered in the model include ACS, knee replacement, stroke and TIA. A one-off disutility is applied for these events in the first cycle in which the event occurs. To account for the impact of living with musculoskeletal disorder, patients receiving knee replacement are assumed to have the disutility applied for three years. Disutility values are taken from Søltoft et al[38] and Sullivan et al.[39]

Sullivan et al compiled a UK-based catalogue of EQ-5D index scores based on 79,522 individuals with completed EQ-5D scores. Utilities from the Søltoft et al. and Sullivan et al. sources are consistent with the NICE reference case. If a fatal event occurs in the acute event state then the fatal event is assumed to occur at the mid-point of the cycle so only half

the acute event disutility is applied and the patient moves to the dead health state. The disutility values associated with acute events are shown in Table 37(CS Table 37).

Disutilities related to bariatric surgery are treated in the model as one-off disutilities applied to the proportion of patients receiving bariatric surgery in each cycle. The disutility represents the decrement in quality of life associated with the surgical procedure and the related complications, based upon Campbell et al[40] and was estimated to be -0.184.

Two adverse events were included in the model for non-severe hypoglycaemia and severe gastrointestinal event. The disutilities for these two adverse events are shown in Table 37(CS Table 40).

4.2.7.3.3 Health state utility values

Health state values for T2D, OSA, Post ACS and Post-stroke are shown in Table 37 (CS Table 38) and use the same sources as the acute events.[38 39] When health states combine two or more obesity complications (e.g. T2D + Post ACS), the utility decrement for this health state is calculated by adding the utility decrements for each of the individual complications. The CS states that Gough et al.[41] concluded that HRQoL decrements associated with T2D and obesity showed no significant interaction and thus could be assumed to be additive. We note that the NICE technical support document 12[42] recommends multiplicative decrements. However, from studies that have reported multiple co-morbidities for diabetes,[39 43] we agree with the company and consider it is reasonable to treat co-morbidities as independent and add utility decrements. In addition, we note that this approach was also taken in TA664.[4]

We note that the decrement for type 2 diabetes is lower than reported in other studies such as Sullivan et al and Ara et al.[44] We use the utility decrement from Sullivan et al[39] (-0.0714) in a scenario in section 6. The decrement for knee replacement is lower in other sources, such as Sullivan et al.[39] (decrement -0.099). We conduct a scenario analysis using the value from Sullivan et al in section 6.

Table 37 Summary of utility values for cost-effectiveness analysis in the company's economic model

State	Utility value: mean (standard error)	Source
Baseline utility	0.901*	Søltoft et al. [38]
ACS	-0.063 (0.046)	Sullivan et al. [39]

Stroke	-0.117 (0.012)	Sullivan et al. [39]
TIA	-0.033 (0.022)	Sullivan et al. [39]
Knee replacement	-0.194 (0.048)**	Søltoft et al. [38]
T2D	-0.037 (0.009)	Søltoft et al. [38]
OSA	-0.038 (0.010)	Søltoft et al. [38]
Post ACS	-0.037 (0.026)	Sullivan et al. [39]
Post-stroke	-0.035 (0.021)	Sullivan et al. [39]
Bariatric surgery	-0.184 (0.046)	Campbell et al[40]
Non-severe hypoglycaemia	-0.0062 (0.002)	Foos et al[45]
Severe gastrointestinal	-0.050 (0.0002)	TA494[12]
Key: ACS, acute coronary syndrome; OSA, obstructive sleep apnoea; T2D, type 2 diabetes; TIA, transient ischemic attack. **Note:** *Baseline for BMI >30 + 1 or more comorbidities is 0.901; Baseline for BMI >35 + non-diabetic hyperglycaemia and high CVD risk is 0.889; Coefficients were not varied; ** Literature value multiplied by three to account for three years of living with osteoarthritis		

Source: CS Table 41

ERG conclusion on HRQoL

The company's approach to estimating utility values is generally reasonable and consistent with the NICE reference case. The ERG notes that the utilities used in this appraisal are largely the same as those used in TA664[4], with the exception of the utility values used for non-severe hypoglycaemia adverse event.

4.2.8 Resources and costs

The costs included in the economic model consist of drug acquisition costs for weight loss treatments, costs for obesity monitoring, health state management costs, acute event costs (including knee replacement and bariatric surgery) and costs for managing AEs.

The company conducted a SLR to identify studies reporting cost and health care resource use data for the treatment of patients with obesity. More details on the review are discussed briefly in this report in section 4.1 and CS Appendix G and I. Three studies were identified but the company reports that the studies did not focus on the patient population or treatments identified as relevant to the decision problem and were therefore not used in the model. The ERG considers that the company's literature review is likely to reflect the available evidence and agrees that the identified studies are not relevant for this appraisal.

The company has conducted a targeted literature review for the costs used in the model. CS Table 47 shows a comparison between the costs used in the current appraisal with those used TA664[4] for liraglutide 3.0 mg.

4.2.8.1 Drug acquisition and monitoring costs

As detailed in section 2.2.2, semaglutide is self-administered once weekly as a subcutaneous injection. The maintenance dose of semaglutide is 2.4 mg (Table 38). The dose is gradually increased over 16 weeks. The titration dose of each of the 16 weeks is shown in CS Table 42. Semaglutide has a list price of [redacted] per pack and each pack contains four pre-filled pens containing a 2.4mg dose.

Liraglutide is administered daily in a similar manner to semaglutide 2.4 mg. The maintenance dose of liraglutide is 3.0 mg and the dose is increased over the first four weeks of treatment. The titration dose of each of the four weeks is shown in CS Table 42. The list price for liraglutide 3.0mg is £196.20 per pack and each pack contains five pre-filled pens containing 18mg of liraglutide (Table 38).

Liraglutide 3.0 mg is available with a confidential price discount (PAS price). Semaglutide 2.4 mg does not currently have an agreed PAS (discussions are ongoing with NHS England). All analyses in this report are for the list price with additional analyses with the PAS prices reported by the ERG in a confidential appendix.

The cost of obesity monitoring includes cost for routine visits, examinations (GP visits and nurse visits and blood tests, see Table 38). The annual monitoring costs is £248.90. The breakdown of the monitoring costs is shown in CS Table 42. There is also an annual cost for blood pressure treatment of £17.66. The ERG considers that monitoring costs have been underestimated as they do not include the costs for dietitian and specialist consultations. In response to the ERG's question B20 in the appraisal for liraglutide 3.0 mg (TA664), the company estimated an annual cost of £353.60. However, we note that as monitoring costs are applied equally across arms, changes to the monitoring costs will have no effect on the ICERs.

Table 38 Treatment and monitoring costs (list price)

Treatment costs	Cost (£)	Description and references
Semaglutide (2.4 mg/ week)	[redacted]	4 pens, 2.4mg per pen Maintenance dose 2.4mg / week
Liraglutide (3.0 mg/ day)	£196.20 per pack	Maintenance period dose: 3.0mg per day:
Monitoring costs for obesity, annual	£248.90	Annual frequency (assumed equal to orlistat and rimonabant) * cost for 3 types of visits:

Treatment costs	Cost (£)	Description and references
		• GP visit: Frequency: 4x10 mins • Nurse visit: Frequency: 8x15 mins • Blood test (1 test)
Blood pressure treatment	17.66	Average annual cost of ACE inhibitor treatment.

Source: CS Table 42

4.2.8.2 Health state costs

The annual health state costs for obesity related complications include the costs for monitoring and treating a given disease and are shown in Table 39 (CS Table 43). In addition to the diabetes health state costs, there is a cost for insulin treatment and oral drugs for diabetes. The costs of health states including multiple obesity complications are calculated by summing the costs associated with each condition. The costs are derived from UK published studies. Costs are updated to 2020 costs using PSSRU[46] inflation indices. The costs of T2D microvascular complications are estimated by summing the costs associated with each condition from the study by Capehorn et al.[47] The ERG notes that costs of T2D microvascular complications are applied from onset of T2D. However, the ERG consider this unlikely as the risk of complications increase with the time since diagnosis.[48] The ERG has estimated the costs reported in Capehorn et al to be lower than used by the company of £507 per year and use these costs in the ERG base case in section 6.

Table 39 Annual obesity related complication costs applied to health states in the model

State costs	Cost (£)	Source
T2D microvascular complications costs	940.86	Capehorn et al[47]
T2D treatment – average of insulin and oral treatments	551.89	Capehorn et al[47]
Non-diabetic hyperglycaemia	54.00	NHS diabetes prevention programme
MI 1st year, excl. acute event cost	1,174.12	Alva et al[49]
Unstable angina 1st year, excl. acute event cost	1,056.18	Alva et al[49]
Post-acute coronary syndrome	846.29	Alva et al [49]
Stroke 1st year, excl. acute event cost	1,333.67	Alva et al[49]
Transient ischaemic attack, 1st year	1,338.77	Danese et al [50]
Post-stroke (stroke and TIA, in year following the event)	944.69	Alva et al[49]
Sleep apnoea cost	1,018.19	NHS reference costs 2018/19 [51]

State costs	Cost (£)	Source
Key: MI, myocardial infarction; T2D, Type 2 diabetes; TIA, transient ischemic attack.		

The sleep apnoea cost is applied to the proportion of patients with sleep apnoea. The cost is taken from NHS reference costs [51]. However, we consider the costs may be overestimated. The long-term annual costs for continuous positive airway pressure machines for sleep apnoea were estimated in Sharples et al[52] as £251.99 per year. We use this cost (inflated to 2020 prices) in the ERG base case in section 6.3.

4.2.8.3 Acute event costs, bariatric surgery and adverse event costs

The model includes one-off costs for the obesity related acute events unstable angina, MI stroke, TIA and knee replacement. These costs include the cost of management, including hospitalisation for the acute event. The cost of knee replacement also includes the cost of pre-surgery visits and examinations and post-surgery follow-up. Costs are taken from NHS reference costs 2018/19 and inflated to 2020 prices. The acute event costs are shown in CS Table 46.

Bariatric surgery is applied as a one-off cost and includes preoperative management, post-operative follow-up and surgery related complications. The average -procedure cost is calculated as the weighted average of the three types of procedure. Bariatric surgery costs are shown in CS Table 45.

The one-off costs for adverse events for non-severe hypoglycaemia and severe gastrointestinal events are £4.09 and £144.01 respectively.

ERG conclusion on resources and costs

The company's approach to resources and costs in the economic model are consistent with the NICE reference case and the previous technology appraisal for liraglutide 3.0 mg (TA664). Some of the cost estimates in TA664 have been updated based on a targeted literature search (CS Table 47). The approach is largely reasonable, with the exception of i) the costs for sleep apnoea and ii) applying the microvascular complication costs from onset of T2D.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The company reported their base case results in CS Table 52, reproduced below in Table 40**Error! Reference source not found.** for the population of people with BMI ≥ 30 kg/m^2 with one or more obesity related co-morbidities. They also conducted a subgroup analysis for the population of people with BMI ≥ 35 kg/m^2, non-diabetic hyperglycaemia and high risk for CVD, comparing the cost-effectiveness of semaglutide 2.4mg versus liraglutide 3.0mg (CS Table 53). The company did not include diet and physical activity as a comparator for this analysis. We present the incremental results for all the comparators below.

The cost-effectiveness results, reproduced below in Table 40 and Table 41, are presented with list prices for all the treatment arms. The results with the PAS price discount for liraglutide 3.0mg is presented in a confidential addendum to this report.

Table 40 Company base case results for semaglutide 2.4 mg versus diet and physical activity (list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Diet and physical activity	[redacted]	17.924	15.269				
Semaglutide 2.4 mg	[redacted]	17.957	15.361	[redacted]	0.034	0.092	[redacted]
Key: ICER, incremental cost-effectiveness ratio; LYG, life year gained; QALY, quality-adjusted life year.							

Source: Reproduced from CS Table 52

Table 41 Subgroup results for semaglutide 2.4 mg versus liraglutide 3.0mg (list price)

Technologies	Total costs (£)	Total LYG	Total QALYs	Incr. costs (£)	Incr. LYG	Incr. QALYs	ICER (£/QALY)
Diet and physical activity	[redacted]	17.288	14.311				
Liraglutide 3.0 mg	[redacted]	17.331	14.401	[redacted]	0.043	0.090	[redacted]
Semaglutide 2.4 mg	[redacted]	17.349	14.444	[redacted]	0.018	0.043	[redacted]
Key: ICER, incremental cost-effectiveness ratio; LYG, life year gained; QALY, quality-adjusted life year.							

Source: Reproduced from CS Table 57

5.2 Company's sensitivity analyses

5.2.1 Deterministic sensitivity analyses

The CS reports results for the one-way, deterministic sensitivity analysis in the tornado plots for: the base case in CS Figure 18 and the liraglutide-eligible subgroup in CS Figure 20 respectively. The ranges of variation for input parameters were based on confidence intervals obtained from standard errors of the mean (where available), or simple assumed percentages where empirical evidence was unavailable. The results indicated that the starting BMI of the cohort, the discount rate for QALYs, and the weight reduction at the start of Year 2 with diet and physical activity have the largest impact on the cost-effectiveness results.

5.2.2 Probabilistic sensitivity analysis

The company conducted a probabilistic sensitivity analysis (PSA), with input parameter distributions as reported in CS Tables 48-50. The company's probabilistic results are reported in CS Table 53 (base case) and Table 58 (liraglutide-eligible subgroup analysis). The cost-effectiveness scatter plot and acceptability curve for the base case are shown in CS Figures 16 and 17 respectively. For the liraglutide-eligible subgroup analysis, they presented the scatter plot in CS Figure 19. The ERG confirms that the probabilistic results for the base case are similar to the deterministic results. The probabilistic ICER for semaglutide 2.4mg compared with diet and physical activity is [redacted] per QALY gained compared with [redacted] per QALY for the deterministic ICER.

5.2.3 Scenario analysis

The company presented fifteen scenario analyses (CS Table 56 for the base case and Table 60 for the liraglutide-eligible subgroup analyses). We reproduce the results of the scenario analyses for the base case and the scenario analyses in Table 42 below.

Table 42: Scenario analyses explored in the economic model (list price)

Scenario	Base case (semaglutide 2.4 mg vs diet and physical activity)	Liraglutide-eligible subgroup (semaglutide 2.4 mg vs liraglutide 3.0mg)
	ICER (£/QALY)	ICER (£/QALY)
Base case	[redacted]	[redacted]
No stopping rule (using treatment policy estimand)	[redacted]	[redacted]
Post hoc analysis efficacy for subgroup (using efficacy data from patients with BMI ≥30 and one or more comorbidities.)	[redacted]	[redacted]
1-year catch up rate	[redacted]	[redacted]
2-year catch up rate	[redacted]	[redacted]
20-year time horizon	[redacted]	[redacted]
30-year time horizon	[redacted]	[redacted]
No bariatric surgery	[redacted]	[redacted]
Bariatric surgery eligibility threshold at BMI 47 kg/m^2	[redacted]	[redacted]
Bariatric surgery incidence at 0.57% per year	[redacted]	[redacted]
Framingham Offspring risk equation for incidence of T2D	[redacted]	[redacted]
Framingham Heart Study risk equation for incidence of first CVD event in NGT	[redacted]	[redacted]
Framingham Recurring Coronary Heart Disease risk equation for incidence of recurrent CVD event in T2D	[redacted]	[redacted]
QRisk3 risk equation for incidence of first CVD event in T2D	[redacted]	[redacted]
Patients with non-diabetic hyperglycaemia do not develop T2D immediately after a CVD event	[redacted]	[redacted]
Alternative baseline utilities, derived as a function of BMI based on SCALE data.	[redacted]	[redacted]

Source: reproduction of CS Table 56

For the base case, the model results are most sensitive to using a one year catch-up rate related to BMI and glycaemic reversal after treatment discontinuation, followed by the scenario with no stopping rule, the two year-catch up rate and using the risk equation using Framingham Offspring risk equation for the incidence of T2D. For these scenarios, the base case ICER [redacted]. Using a bariatric surgery eligibility threshold of BMI 47kg/m^2 as well the scenario with no bariatric surgery [redacted].

For the liraglutide-eligible subgroup analyses, semaglutide 2.4 mg remained [redacted].

5.3 Model validation and face validity check

The company approach to validation is described in CS section B.3.10. For quality control, the model checks included conducting top-down tests (i.e. changing input parameters), model functionality (i.e. testing all key parameters, extreme value tests), and internal consistency (i.e. accuracy of input data against source data). For external validation, the company cited the publication by Lopes et al [53] that reported that the model predicted CVD and T2D with a good degree of accuracy. They did not provide any information on internal validity, i.e. comparing the model results with outputs from the phase 3 STEP 1 trial or the SCALE 1839 trial.

The ERG conducted a series of quality checks of the company model. These included: checking that the input parameters in the model matched the values in the CS and in the original sources; and validating the results of the scenario and sensitivity analyses as reported by the company. We also conducted a series of 'white box' and 'black box' checks to validate the model. We did not identify any errors in the model. However, the ERG were unable to replicate the scenario using alternative baseline utilities, derived as a function of BMI based on SCALE data.

5.3.1 Internal validation

For internal validation, the company provided a comparison of the modelled estimated clinical events (mean BMI, SBP and total cholesterol) for the first two years with the trial product estimand data, as response to clarification question B10. We note that the change in the clinical outcomes for semaglutide 2.4mg at 2 years from baseline are slightly higher compared to that of the change in the trial outcomes at 68 weeks from baseline. However, we do not anticipate these differences to impact the overall cost-effectiveness results.

Table 43: Comparison of modelled outcomes versus STEP 1 trial reported outcomes

Parameter	Modelled outcome change at 2 years from baseline	Trial outcome change at 68 weeks from baseline	Modelled outcome change at 2 years from baseline	Trial outcome change at 68 weeks from baseline
	Semaglutide 2.4 mg		Diet and physical activity	
BMI (kg/m^2)	-6.37	-6.27	-0.92	-0.95

Parameter	Modelled outcome change at 2 years from baseline	Trial outcome change at 68 weeks from baseline	Modelled outcome change at 2 years from baseline	Trial outcome change at 68 weeks from baseline
SBP (mmHg)	-7.08	-7.08	-1.14	-1.14
Total cholesterol (mg/dL)	-8.44	-7.45	0.18	-1.49
Key: BMI, body mass index; SBP, systolic blood pressure.				

Source: Reproduced from Tables 6 and 7 from the company's response to ERG clarification B10.

5.3.2 External validation

In their response to clarification question B11, the company compared the costs and QALYs associated with the comparator arm – diet and physical activity – of the current appraisal with previous TA664 for the subgroup with BMI ≥ 35kg/m^2 and non-diabetic hyperglycaemia and high risk of CVD, reproduced below in Table 44. They argued the difference in results were attributed to the inclusion of mortality adjusted for BMI, updated costs and the inclusion of the provision for weight to return to the value of natural progression at the end of the catch-up period.

Table 44 Comparison of results from current appraisal and TA664

Technologies	Costs (£)	QALYs
Diet & physical activity (SCALE-TA664)	£19,992	15.18
Diet & physical activity (STEP 1)	£28,371	14.31
Key: QALY, quality-adjusted life year.		

Source: Reproduced from Table 8 from company's response to clarification question B11

The company used the patient characteristics and efficacy data for diet and physical activity from the SCALE 1839 trial in the current appraisal, results are reproduced in Table 45 below. We agree with the company's conclusion that the results between the two appraisals are similar when adjusted for patient characteristics and efficacy data.

Table 45 Comparison of results from current appraisal and TA664 (adjusted for patient characteristics and efficacy data)

Technologies	Costs (£)	QALYs
Diet & physical activity (SCALE-TA664)	£27,597	14.60
Diet & physical activity (STEP 1)	£28,371	14.31
Key: QALY, quality-adjusted life year.		

Source: Reproduced from Table 9 from company's response to clarification question B11

6 ERG'S ADDITIONAL ANALYSES

6.1 Corrections to the company's base case

The ERG did not identify any errors that affected the company's base case analysis. However, we conducted further scenarios for those aspects where we considered uncertainties remained. These are discussed in section 6.2.

6.2 Impact on the ICER of additional analyses undertaken by the ERG

The ERG conducted a series of scenarios on the company's base case and liraglutide-eligible subgroup analyses. These are listed in Table 46 and Table 49 below respectively. We note that change in mean starting BMI has the most significant impact on the cost-effectiveness results. Across the scenarios conducted for the company's base case, the ICERs for semaglutide 2.4mg vs diet and physical activity vary between [redacted] (Scenario: Mean BMI of 42.5kg/m^2) and [redacted] (Scenario: Mean BMI of 32.5 kg/m^2).

Table 46 Scenarios conducted by the ERG on the company's base case

Assumption	Treatments	Total costs	Total QALYs	ICER (£/QALY)
Company base-case	Diet & physical activity	[redacted]	15.269	[redacted]
	Semaglutide 2.4mg		15.361	
Using the patient characteristics for all patients in STEP 1	Diet & physical activity	[redacted]	15.943	[redacted]
	Semaglutide 2.4mg	[redacted]	16.041	
Using the subgroup efficacy for those with BMI >=30 and one or more comorbidities.	Diet & physical activity	[redacted]	15.273	[redacted]
	Semaglutide 2.4mg	[redacted]	15.367	
Yearly increase in weight of 0.296kg (Iyen et al.)	Diet & physical activity	[redacted]	15.422	[redacted]
	Semaglutide 2.4mg	[redacted]	15.521	
Age at which weight no longer increases: 66 years	Diet & physical activity	[redacted]	15.276	[redacted]
	Semaglutide 2.4mg	[redacted]	15.369	
Using QRISK3 for incidence of first CVD event in T2D	Diet & physical activity	[redacted]	15.136	[redacted]
	Semaglutide 2.4mg	[redacted]	15.235	
Disutility for T2D: -0.0714 (Sullivan et al.)	Diet & physical activity	[redacted]	15.138	[redacted]
	Semaglutide 2.4mg	[redacted]	15.240	
Disutility for knee replacement: -0.099 (Sullivan et al.)	Diet & physical activity	[redacted]	15.323	[redacted]
	Semaglutide 2.4mg	[redacted]	15.414	
Mean starting BMI of 32.5 kg/m^2	Diet & physical activity	[redacted]	16.453	[redacted]
	Semaglutide 2.4mg	[redacted]	16.533	
Mean starting BMI of 37.5 kg/m^2	Diet & physical activity	[redacted]	15.510	[redacted]
	Semaglutide 2.4mg	[redacted]	15.615	
Mean starting BMI of 42.5 kg/m^2	Diet & physical activity	[redacted]	14.495	[redacted]
	Semaglutide 2.4mg	[redacted]	14.617	

For the liraglutide-eligible subgroup, the scenario with Mean starting BMI of 37.5 kg/m^2 has the most significant impact on the model results with the ICER for liraglutide 3.0mg vs diet and physical activity increasing from £[redacted] to £[redacted].

Table 47 Scenarios conducted by the ERG on the company's liraglutide-eligible subgroup

Assumption	Treatments	Total costs	Total QALYs	ICER (£/QALY)
Company base-case	Diet & physical activity	[redacted]	14.311	
	Liraglutide 3.0mg	[redacted]	14.401	[redacted]
	Semaglutide 2.4mg	[redacted]	14.444	[redacted]
Using efficacy data for those with BMI >=35 and one or more comorbidities.	Diet & physical activity	[redacted]	14.420	
	Liraglutide 3.0mg	[redacted]	14.509	[redacted]
	Semaglutide 2.4mg	[redacted]	14.543	[redacted]
Yearly increase in weight of 0.296kg (Iyen et al.)	Diet & physical activity	[redacted]	14.450	
	Liraglutide 3.0mg	[redacted]	14.540	[redacted]
	Semaglutide 2.4mg	[redacted]	14.583	[redacted]
Age at which weight no longer increases: 66 years	Diet & physical activity	[redacted]	14.380	
	Liraglutide 3.0mg	[redacted]	14.470	[redacted]
	Semaglutide 2.4mg	[redacted]	14.513	[redacted]
Using QRISK3 for incidence of first CVD event in T2D	Diet & physical activity	[redacted]	14.106	
	Liraglutide 3.0mg	[redacted]	14.206	[redacted]
	Semaglutide 2.4mg	[redacted]	14.251	[redacted]
Disutility for T2D: -0.0714 (Sullivan et al.)	Diet & physical activity	[redacted]	14.114	
	Liraglutide 3.0mg	[redacted]	14.216	[redacted]
	Semaglutide 2.4mg	[redacted]	14.261	[redacted]
Disutility for knee replacement: -0.099 (Sullivan et al.)	Diet & physical activity	[redacted]	14.380	
	Liraglutide 3.0mg	[redacted]	14.469	[redacted]
	Semaglutide 2.4mg	[redacted]	14.511	[redacted]
Mean starting BMI of 37.5 kg/m^2	Diet & physical activity	[redacted]	15.260	
	Liraglutide 3.0mg	[redacted]	15.334	[redacted]
	Semaglutide 2.4mg	[redacted]	15.377	[redacted]

6.3 ERG's preferred assumptions

The ERG's preferred assumptions are listed in the Table 48.

Table 48 ERG preferred assumptions

Model aspect	Company's assumption	ERG preferred assumption
Transition of patients with non-hyperglycaemia	Patients develop T2D immediately after a CVD event.	The NCE committee from TA664 concluded that there was no good evidence to determine the proportion of people who develop type 2 diabetes after a cardiovascular event. Furthermore, clinical advice to the ERG suggest that it is not possible to assume that patients will develop T2D

		after a CVD event. Further details in section 4.2.2.
Body weight	Patients undergo a natural increase of weight per year of 0.463kg. This is based on an increase in BMI of 0.1447 kg/m^2 for males and 0.1747 kg/m^2 for females per year for the diet and physical activity arm from the study by Ara et al.[24]	Patients undergo a natural increase of weight per year of 0.296 kg. This is based on a more recent study by Iyen et al.[25] that included a cohort of 264,230 individuals in the UK. The study estimated a mean increase in BMI of 1.06 kg/m^2 per 10 years. Further details in section 4.2.2**Error! Reference source not found.**.
Maximum age of weight increase	68 years	66 years.
Direction of weight change at maximum age	Weight remains constant after the maximum age of 68 years.	There is a steep decrease in individuals' BMI based on the study by Zaninotto et al[26] that explored BMI trajectories in the English Longitudinal Study of Ageing. We assume that the weight decrease post maximum age is similar to that of weight increase before reaching maximum age of weight gain. Further details in section 4.2.2.**Error! Reference source not found.**
Costs of microvascular complication and sleep apnoea	Microvascular complication: £941 Sleep apnoea: £1018	Microvascular complication: £398[47] Sleep apnoea: £274[52]

The cumulative effect of the ERG's preferred assumptions to the company's analyses are shown in Table 49 and Table 50. Applying the ERG preferred assumptions increases the company's base case ICER for semaglutide 2.4mg versus diet and physical activity from [redacted] to [redacted] per QALY. For the liraglutide-eligible subgroup, while the ICER for liraglutide 3.0mg versus diet and physical activity increases from [redacted] to [redacted] semaglutide 2.4mg [redacted] compared to liraglutide 3.0mg.

Table 49 Cumulative change from the company base case to ERG base case with ERG's preferred assumptions

Assumption	Treatments	Total costs	Total QALYs	ICER (£/QALY)
Company base-case	Diet & physical activity	[redacted]	15.269	[redacted]
	Semaglutide 2.4mg	[redacted]	15.361	
Patients with pre-diabetes do not transition to T2D after CVD events	Diet & physical activity	[redacted]	15.329	[redacted]
	Semaglutide 2.4mg	[redacted]	15.419	
	Diet & physical activity	[redacted]	15.484	[redacted]

+ Mean increase of weight by 0.296 kg per year	Semaglutide 2.4mg	[redacted]	15.582	
+ Mean decrease in weight after age 66 years: 0.296 kg per year	Diet & physical activity	[redacted]	15.540	[redacted]
	Semaglutide 2.4mg	[redacted]	15.634	
+ Age at which weight no longer increases: 66 years	Diet & physical activity	[redacted]	15.562	[redacted]
	Semaglutide 2.4mg	[redacted]	15.656	
+ Annual cost of microvascular complication, £398	Diet & physical activity	[redacted]	15.562	[redacted]
	Semaglutide 2.4mg	[redacted]	15.656	
+ Annual cost of sleep apnoea, £274	Diet & physical activity	[redacted]	15.562	[redacted]
	Semaglutide 2.4mg	[redacted]	15.656	
ERG base case	Diet & physical activity	[redacted]	15.562	[redacted]
	Semaglutide 2.4mg	[redacted]	15.656	

Table 50 Cumulative change from company liraglutide-eligible subgroup results to the ERG liraglutide-eligible subgroup results with the ERG's preferred assumptions

Assumption	Treatments	Total costs	Total QALYs	Incremental ICER (£/QALY)
Company base-case	Diet & physical activity	[redacted]	14.311	
	Liraglutide 3.0mg	[redacted]	14.401	[redacted]
	Semaglutide 2.4mg	[redacted]	14.444	[redacted]
Patients with pre-diabetes do not transition to T2D after CVD events	Diet & physical activity	[redacted]	14.419	
	Liraglutide 3.0mg	[redacted]	14.505	[redacted]
	Semaglutide 2.4mg	[redacted]	14.548	[redacted]
+ Mean increase of weight by 0.296 kg per year	Diet & physical activity	[redacted]	14.562	
	Liraglutide 3.0mg	[redacted]	14.648	[redacted]
	Semaglutide 2.4mg	[redacted]	14.690	[redacted]
+ Mean decrease in weight after age 66 years: 0.296 kg per year	Diet & physical activity	[redacted]	14.642	
	Liraglutide 3.0mg	[redacted]	14.727	[redacted]
	Semaglutide 2.4mg	[redacted]	14.770	[redacted]
+ Age at which weight no longer increases: 66 years	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]
+ Annual cost of microvascular complication, £398	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]
+ Annual cost of sleep apnoea, £274	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]
ERG base case	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]

The ERG conducted a series of scenario analyses on the base case and the liraglutide-eligible subgroup, shown below in Table 51 and Table 52. For the base case, the ICER for semaglutide 2.4mg versus diet and physical activity varied between [redacted] (Scenario: Mean starting BMI of 42.5 kg/m^2) and [redacted] (Scenario: catch up rate of 1 year). For the liraglutide-eligible subgroup, semaglutide 2.4mg was [redacted] compared to liraglutide 3.0mg for all scenarios.

Table 51 Scenarios conducted on the ERG base case

Assumption	Treatments	Total costs	Total QALYs	ICER (£/QALY)
ERG base-case	Diet & physical activity	[redacted]	15.562	[redacted]
	Semaglutide 2.4mg	[redacted]	15.656	
Mean starting BMI of 32.5 kg/m^2	Diet & physical activity	[redacted]	16.762	[redacted]
	Semaglutide 2.4mg	[redacted]	16.839	
Mean starting BMI of 37.5 kg/m^2	Diet & physical activity	[redacted]	15.766	[redacted]
	Semaglutide 2.4mg	[redacted]	15.870	
Mean starting BMI of 42.5 kg/m^2	Diet & physical activity	[redacted]	14.775	[redacted]
	Semaglutide 2.4mg	[redacted]	14.895	
No stopping rule	Diet & physical activity	[redacted]	15.569	[redacted]
	Semaglutide 2.4mg	[redacted]	15.651	
Catch-up rate: 1 year	Diet & physical activity	[redacted]	15.541	[redacted]
	Semaglutide 2.4mg	[redacted]	15.609	
Catch-up: 2 years	Diet & physical activity	[redacted]	15.557	[redacted]
	Semaglutide 2.4mg	[redacted]	15.634	
Catch-up: 4 years	Diet & physical activity	[redacted]	15.578	[redacted]
	Semaglutide 2.4mg	[redacted]	15.685	
Treatment duration: 3 years	Diet & physical activity	[redacted]	15.563	[redacted]
	Semaglutide 2.4mg	[redacted]	15.693	
Using QRISK3 for incidence of first CVD event in T2D	Diet & physical activity	[redacted]	15.423	[redacted]
	Semaglutide 2.4mg	[redacted]	15.524	

Table 52 Scenarios conducted on the ERG liraglutide-eligible subgroup

Assumption	Treatments	Total costs	Total QALYs	Incremental ICER (£/QALY)
ERG base-case	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.745	[redacted]
	Semaglutide 2.4mg	[redacted]	14.788	[redacted]
Mean starting BMI of 37.5 kg/m^2	Diet & physical activity	[redacted]	15.580	
	Liraglutide 3.0mg	[redacted]	15.651	[redacted]
	Semaglutide 2.4mg	[redacted]	15.694	[redacted]
Mean starting BMI of 42.5 kg/m^2	Diet & physical activity	[redacted]	14.596	
	Liraglutide 3.0mg	[redacted]	14.672	[redacted]
	Semaglutide 2.4mg	[redacted]	14.726	[redacted]
No stopping rule	Diet & physical activity	[redacted]	[redacted]	
	Liraglutide 3.0mg	[redacted]	14.743	[redacted]
	Semaglutide 2.4mg	[redacted]	14.802	[redacted]
Catch-up rate: 1 year	Diet & physical activity	[redacted]	14.638	
	Liraglutide 3.0mg	[redacted]	14.694	[redacted]
	Semaglutide 2.4mg	[redacted]	14.718	[redacted]
Catch-up rate: 2 years	Diet & physical activity	[redacted]	14.649	
	Liraglutide 3.0mg	[redacted]	14.722	[redacted]
	Semaglutide 2.4mg	[redacted]	14.760	[redacted]
Catch-up rate: 4 years	Diet & physical activity	[redacted]	14.667	
	Liraglutide 3.0mg	[redacted]	14.765	[redacted]
	Semaglutide 2.4mg	[redacted]	14.814	[redacted]
Treatment duration: 3 years	Diet & physical activity	[redacted]	14.659	
	Liraglutide 3.0mg	[redacted]	14.768	[redacted]
	Semaglutide 2.4mg	[redacted]	14.830	[redacted]
	Diet & physical activity	[redacted]	14.439	

Using QRISK3 for incidence of first CVD event in T2D	Liraglutide 3.0mg	[redacted]	14.535	[redacted]
	Semaglutide 2.4mg	[redacted]	14.580	[redacted]

6.4 Conclusions on the cost effectiveness evidence

The company developed a de novo model, based on the model developed for the NICE technology appraisal TA664 for liraglutide 3.0 mg for managing overweight and obesity.[4] The ERG considers the model structure is appropriate to reflect this condition and the treatment pathway and is consistent with the NICE reference case.

The model uses intermediate clinical outcomes to extrapolate to morbidity events and mortality beyond the trial period. As such, there is inherent uncertainty in the cost-effectiveness results from the modelling. However, it is reassuring that the company has provided validation of the extrapolation of clinical outcomes.

The company base case ICER for semaglutide 2.4mg vs diet and physical activity is [redacted] per QALY. The results are most sensitive to changes to the starting BMI of the cohort, the catch-up rate (time for patients to regain weight) and the incorporation of the stopping rule for non-responders. Semaglutide 2.4mg is more cost-effective for those with higher BMI, therefore in the base case cost-effectiveness estimate, those with higher BMI are compensating for those with lower BMI. For this reason, the ERG presents the results for different BMI ranges. The catch-up rate is uncertain as no follow-up data were available for weight gain in the three years after stopping treatment. However, this duration of catch-up has previously been accepted by the NICE committee in TA664. There is some uncertainty around the inclusion or the stopping rule for non-responders as it was not included in the company's clinical trial and it is still unclear whether the marketing authorisation for semaglutide 2.4mg will include it.

The ERG suggests the following changes to the parameters and assumptions used in the company model:

i) Patients with pre-diabetes do not transition to T2D after CVD events,
ii) Alternative natural history increase in population's weight over time,
iii) Reduced annual cost of microvascular complications for T2D and sleep apnoea.

The ERG base case including these changes is [redacted] per QALY for semaglutide 2.4mg versus diet and physical activity.

7 END OF LIFE

Semaglutide 2.4mg is not suitable to be considered as an end-of-life treatment as the population to be treated with it does not fulfil the criteria to have an expected life expectancy of less than 24 months.

8 References

1. National Institute for Health and Care Excellence. Obesity: identification, assessment and management [CG189] 2014 [updated 27 November 2014. Available from: https://www.nice.org.uk/guidance/cg189/resources/obesity-identification-assessment-and-management-pdf-35109821097925 accessed 6 September 2021.

2. Davies M, Færch L, Jeppesen OK, et al. Semaglutide 2·4 mg once a week in adults with overweight or obesity, and type 2 diabetes (STEP 2): a randomised, doubleblind, double-dummy, placebo-controlled, phase 3 trial. *Lancet* 2021;397:971-84.

3. Wadden TA, Bailey TS, Billings LK, et al. Effect of Subcutaneous Semaglutide vs Placebo as an Adjunct to Intensive Behavioral Therapy on Body Weight in Adults With Overweight or Obesity: The STEP 3 Randomized Clinical Trial. *JAMA* 2021;325(14):1403-13. doi: 10.1001/jama.2021.1831

4. National Institute for Health and Care Excellence. Liraglutide for managing overweight and obesity [TA664] 2020 [updated 9 December 2020. Available from: https://www.nice.org.uk/guidance/ta664/resources/liraglutide-for-managing-overweight-and-obesity-pdf-82609259121349 accessed 6 September 2021.

5. Wilding JPH, Batterham RL, Calanna S, et al. Once-Weekly Semaglutide in Adults with Overweight or Obesity. *N Engl J Med* 2021;384(11):989. doi: 10.1056/NEJMoa2032183

6. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. *J Gen Intern Med* 2001;16(9):606-13.

7. Jennings A, Hughes CA, Kumaravel B, et al. Evaluation of a multidisciplinary Tier 3 weight management service for adults with morbid obesity, or obesity and comorbidities, based in primary care. *Clinical Obesity* 2014;4(5):254-66.

8. European Medicines Agency. Summary of Product Characteristics: Saxenda. European Medicines Agency, 2015.

9. Blackshaw J, Montel S, King S, et al. Report of the working group into: Joined up clinical pathways for obesity: Public Health England, 2014.

10. British Obesity and Metabolic Surgery Society. Commissioning guide: Weight assessment and management clinics (tier 3): British Obesity and Metabolic Surgery Society, 2014.

11. National Institute for Health and Care Excellence. Obesity: clinical assessment and management (QS127) 2016 [updated May 2020. Available from: https://www.nice.org.uk/guidance/qs127/resources/obesity-clinical-assessment-and-management-pdf-75545363615173 accessed 6 September 2021.

12. National Institute for Health and Care Excellence. Naltrexone-bupropion for managing overweight and obesity [TA494] 2017 [updated 12 December 2017. Available from: https://www.nice.org.uk/guidance/ta494 accessed 6 September 2021.

13. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. York Publishing Services Ltd.: CRD, 2009.

14. Novo Nordisk. Effect and safety of semaglutide 2.4 mg once-weekly in subjects with overweight or obesity. Clinical Study Report, 2020.

15. Novo Nordisk. IPD-ITC report semaglutide 2.4 mg versus liraglutide 3.0 mg, 2021.

16. National Institute for Health and Care Excellence. Clinical Knowledge Summary - Obesity. Scenario: management: NICE, 2017.

17. Min T, Bain SC. Estimands in diabetes clinical trials. *The Lancet: Diabetes and endocrinology* 2020;8(3):181-83.

18. Aroda VR, Saugstrup T, Buse JB, et al. Incorporating and interpreting regulatory guidance on estimands in diabetes clinical trials: The PIONEER 1 randomized clinical trial as an example. *Diabetes Obes Metab* 2019;21(10):2203-10. doi: 10.1111/dom.13804

19. Pi-Sunyer X, Astrup A, Fujioka K, et al. A Randomized, Controlled Trial of 3.0 mg of Liraglutide in Weight Management. *N Engl J Med* 2015;373(1):11-22. doi: 10.1056/NEJMoa1411892

20. Carlsson Petri KC, Ingwersen SH, Flint A, et al. Once-Weekly in Type 2 Diabetes: A Population Pharmacokinetic Analysis. *Diabetes Therapy* 2018;9(4):1533-47.

21. Overgaard RV, Petri KC, Jacobsen LV, et al. Liraglutide 3.0 mg for Weight Management: A Population Pharmacokinetic Analysis. *Clinical Pharmacokinetics* 2016;55(11):1413-22.

22. Bays H, Pi-Sunyer X, Hemmingsson JU, et al. Liraglutide 3.0 mg for weight management: weight-loss dependent and independent effects. *Current Medical Research and Opinion* 2017;33(2):225-29.

23. Wilding JP, Overgaard RV, Jacobsen LV, et al. Exposure-response analyses of liraglutide 3.0 mg for weight management. *Diabetes, Obesity and Metabolism* 2016;18(5):491-99.

24. 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):3–14. doi: 10.3310/hta16050

25. Iyen B, Weng S, Vinogradova Y, et al. Long-term body mass index changes in overweight and obese adults and the risk of heart failure, cardiovascular disease and mortality: a cohort study of over 260,000 adults in the UK. *BMC Public Health* 2021;21(1) doi: 10.1186/s12889-021-10606-1

26. Zaninotto P, Lassale C. Socioeconomic trajectories of body mass index and waist circumference: results from the English Longitudinal Study of Ageing. *BMJ Open* 2019;9(4):e025309. doi: 10.1136/bmjopen-2018-025309 [published Online First: 2019/04/22]

27. European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials 2020 [updated 17 February 2020. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf accessed 6 September 2021.

28. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017;357:j2099. doi: 10.1136/bmj.j2099 [published Online First: 2017/05/26]

29. Hippisley-Cox J, Coupland C. Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. *BMJ* 2017;359 doi: 10.1136/bmj.j5019

30. Hayes A, Leal J, Gray A, et al. UKPDS outcomes model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82. *Diabetologia* 2013;56(9):1925-33.

31. Wilson PW, Meigs JB, Sullivan L, et al. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. *Arch Intern Med* 2007;167(10):1068-74. doi: 10.1001/archinte.167.10.1068 [published Online First: 2007/05/30]

32. D'Agostino RB, Russell MW, Huse DM, et al. Primary and subsequent coronary risk appraisal: new results from the Framingham study. *Am Heart J* 2000;139(2 Pt 1):272-81. doi: 10.1067/mhj.2000.96469 [published Online First: 2000/01/29]

33. Young T, Shahar E, Nieto FJ, et al. Predictors of sleep-disordered breathing in community-dwelling adults: the Sleep Heart Health Study. *Arch Intern Med* 2002;162(8):893-900. doi: 10.1001/archinte.162.8.893 [published Online First: 2002/04/23]

34. Wendelboe AM, Hegmann KT, Biggs JJ, et al. Relationships between body mass indices and surgical replacements of knee and hip joints. *Am J Prev Med* 2003;25(4):290-5. doi: 10.1016/s0749-3797(03)00218-6 [published Online First: 2003/10/29]

35. Office for National Statistics. National life tables: UK 2020 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables/current.

36. Bhaskaran K, dos-Santos-Silva I, Leon DA, et al. Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3·6 million adults in the UK. *Lancet Diabetes Endocrinol* 2018;6(12):944-53. doi: 10.1016/s2213-8587(18)30288-2

37. Mulnier HE, Seaman HE, Raleigh VS, et al. Mortality in people with type 2 diabetes in the UK. *Diabet Med* 2006;23(5):516-21. doi: 10.1111/j.1464-5491.2006.01838.x [published Online First: 2006/05/10]

38. Søltoft F, Hammer M, Kragh N. The association of body mass index and health-related quality of life in the general population: data from the 2003 Health Survey of England. *Qual Life Res* 2009;18(10):1293-99.

39. Sullivan P, Slejko J, Sculpher M, et al. Catalogue of EQ-5D scores for the United Kingdom. *Medical Decision Making* 2011;31(6):800-04.

40. Campbell J, McGarry L, Shikora S, et al. Cost-effectiveness of laparoscopic gastric banding and bypass for morbid obesity. *Am J Manag Care* 2010;16(7):e174-e87.

41. Gough SC, Kragh N, Ploug UJ, et al. Impact of obesity and type 2 diabetes on health-related quality of life in the general population in England. *Diabetes Metab Syndr Obes* 2009;2:179-84. doi: 10.2147/dmsott.s7088 [published Online First: 2009/01/01]

42. National Institute for Health and Care Excellence. NICE DSU Technicala Support Document 12: The use of health state utility values in decision models http://nicedsu.org.uk/technical-support-documents/utilities-tsd-series/2011 [Available from: http://www.nicedsu.org.uk/wp-content/uploads/2017/05/Population-adjustment-TSD-FINAL.pdf.

43. Hayes A, Arima H, Woodward M, et al. Changes in Quality of Life Associated with Complications of Diabetes: Results from the ADVANCE Study. *Value Health* 2016;19(1):36-41. doi: 10.1016/j.jval.2015.10.010 [published Online First: 2016/01/23]

44. 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. doi: 10.1016/j.jval.2010.10.029 [published Online First: 2011/06/15]

45. Foos V, McEwan P. Conversion of Hypoglycemia Utility Decrements from Categorical Units Reflecting Event History into Event Specific Disutility Scores Applicable to Diabetes Decision Models. *Value in Health* 2018;21:S223. doi: 10.1016/j.jval.2018.04.1506

46. Curtis L, Burns A. Unit Costs of Health and Social Care 2020 2020 [updated 13 Januray 2021. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020 accessed 6 September 2021.

47. Capehorn M, Hallén N, Baker-Knight J, et al. Evaluating the Cost-Effectiveness of Once-Weekly Semaglutide 1 mg Versus Empagliflozin 25 mg for Treatment of Patients with Type 2 Diabetes in the UK Setting. *Diabetes Therapy* 2021;12(2):537-55.

48. Nanayakkara N, Ranasinha S, Gadowski A, et al. Age, age at diagnosis and diabetes duration are all associated with vascular complications in type 2 diabetes. *J Diabetes Complications* 2018;32(3):279-90. doi: 10.1016/j.jdiacomp.2017.11.009 [published Online First: 2018/01/21]

49. Alva M, Gray A, Mihaylova B, et al. The impact of diabetes-related complications on healthcare costs: new results from the UKPDS (UKPDS 84). *Diabetic Medicine* 2015;32(4):459-66.

50. Danese MD, Gleeson M, Kutikova L, et al. Estimating the economic burden of cardiovascular events in patients receiving lipid-modifying therapy in the UK. *BMJ open* 2016;6(8):e011805.

51. National Health Service. National schedule of reference costs: 2018/19 2020 [updated 19 February 2020. Available from: https://webarchive.nationalarchives.gov.uk/20200501111101/https://improvement.nhs.uk/resources/national-cost-collection/ accessed 6 September 2021.

52. Sharples L, Glover M, Clutterbuck-James A, et al. Clinical effectiveness and cost-effectiveness results from the randomised controlled Trial of Oral Mandibular Advancement Devices for Obstructive sleep apnoea-hypopnoea (TOMADO) and long-term economic analysis of oral devices and continuous positive airway pressure. *Health Technol Assess* 2014;18(67):1-296. doi: 10.3310/hta18670 [published Online First: 2014/11/02]

53. Lopes S, Meincke HH, Lamotte M, et al. A novel decision model to predict the impact of weight management interventions: The Core Obesity Model. *Obesity Science & Practice* 2021;7(3):269-80. doi: https://doi.org/10.1002/osp4.495

9 Appendices

9.1 Company and ERG risk of bias assessments for the STEP 1 and SCALE 1839 trials

Appendix 9.1 Table 1A Company and ERG risk of bias assessments for the STEP 1 trial

Criterion	Company judgement	ERG judgement
Was randomisation carried out appropriately?	**Yes** Performed using an interactive web-based response system	**Yes (=low risk of selection bias)** Centrally randomised…. using an IWRS [interactive web-based response system] (CSR section 9.4.2.1)
Was the concealment of treatment allocation adequate?	**Yes** No rationale reported	**Yes (=low risk of selection bias)** Interactive voice or interactive web-based response system would have concealed allocation
Were the groups similar at the outset of the study in terms of prognostic factors?	**Yes** There were no noteworthy differences in baseline characteristics or medical history	**Yes (=low risk of selection bias)** Baseline characteristics were similar in the two treatment groups (CSR Tables 10-2 and 10-3)
Were the care providers, participants, and outcome assessors blind to treatment allocation?	The company did not report an assessment – **assumed "yes"** by ERG Participants & investigators were masked to treatment allocation during the entire trial	**Yes (=low risk of performance and detection biases)** Treatment allocation remained blinded to the subjects, the investigators and to Novo Nordisk during the entire treatment and follow-up period in the main phase of the trial and until after DBL for the main phase of the trial. Semaglutide and diet and physical activity were identical in appearance and were packed and labelled to fulfil the requirements for double-blind procedures (CSR section 9.4.2.2)
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	The company did not report an assessment – **assumed "yes" by ERG**	**Yes (=unclear risk of attrition bias in relation to this aspect of imbalances in missing data)**

	A greater proportion of participants in the semaglutide 2.4 mg group withdrew due to AEs (92 [7%] of 1306 participants) than did in the diet and physical activity group (21 [3.2%] of 655 participants).	A total of 66 patients (5.1%) in the semaglutide 2.4 mg arm and 46 patients (7.0%) in the diet and physical activity arm were withdrawn or withdrew from the STEP 1 study. Differences in the proportion of missing data between trial arms were small and reasons for data missing similar for the two arms (CS Appendix Figure 3) However, in the liraglutide eligible subgroup, participants with missing data had a lower age, and rates of cardiovascular disease, dyslipidaemia, and hypertension than those without missing data (clarification response attachment E). The ERG are unclear whether this would be a source of attrition bias (see section 3.3.5).
Is there any evidence to suggest that the authors measured more outcomes than they reported?	**No**[a] There is no evidence to suggest that the authors measured more outcomes than they reported for the STEP 1 study (clarification response A4)	**No (=low risk of reporting bias)** Protocol specified outcomes were checked by the ERG against the CSR and trial publication. All primary and secondary outcomes were reported
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** Details of the imputation methods for missing data are given in CS Appendix Table 9.[b]	**Yes (=low risk of attrition bias in relation to this aspect of imbalances in missing data)** The ITT principle was followed in the trial (CSR, section 9.7.1.1), and the imputation methods appear generally appropriate.

[a] Reported in clarification response A4.
[b] The methods used to account for missing data in the ITC have been transposed for the SCALE 1839 trial and STEP 1 trial in Appendix Table 9 compared to the descriptions in the ITC report.
Source: partly reproduced from CS Appendix Table 9.

Appendix 9.1 Table 1B Company and ERG risk of bias assessments for the SCALE 1839 trial

Criterion	Company judgement	ERG judgement
Was randomisation carried out appropriately?	**Yes** Performed using a funder-provided telephone or web-based system	**Yes (=low risk of selection bias)** Randomization was performed with the use of a telephone or Web-based system provided (trial publication)
Was the concealment of treatment allocation adequate?	**Yes** No rationale reported	**Yes (=low risk of selection bias)** Interactive voice or interactive web-based response system would have concealed allocation
Were the groups similar at the outset of the study in terms of prognostic factors?	**Yes** There were no noteworthy differences in baseline characteristics or medical history.	**Yes (=low risk of selection bias)** Baseline characteristics were similar in the two groups (trial publication)
Were the care providers, participants, and outcome assessors blind to treatment allocation?	**Yes** Participants & investigators were masked to treatment allocation during the entire trial	**Yes (=low risk of performance and detection biases)** Participants and investigators were masked to treatment allocation during the entire trial…. and visually identical devices were used for subcutaneous injection (Le Roux et al., 2017, page 2)
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	**Yes** A greater proportion of participants in the liraglutide group withdrew due to AEs (199 [13%] of 1501 participants) than did in the diet and physical activity group (46 [6%] of 747)	**Yes (=unclear risk of attrition bias in relation to this aspect of imbalances in missing data)** A total of 1789 patients (71.9%) in the liraglutide group, as compared with 801 patients (64.4%) in the diet and physical activity group, completed 56 weeks of treatment. Differences in the proportion of missing data between trial arms were small and reasons for data missing similar for the two arms (trial publication). However, in the liraglutide eligible subgroup, participants with missing data had a lower

		age, weight and rates of dyslipidaemia, and hypertension than those without missing data (Clarification response attachment E). The ERG are unclear whether this would be a source of attrition bias (see section 3.3.5).
Is there any evidence to suggest that the authors measured more outcomes than they reported?	**No** No rationale reported	**No (=low risk of reporting bias)** Protocol specified outcomes were checked by the ERG against the trial publications (Pi-Sunyer et al., 2015 and Le Roux et al., 2017). All primary and secondary outcomes were reported.
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	**No** The pre-specified efficacy analyses used data from the full analysis set of all randomised individuals who received at least one treatment dose and had at least one post-baseline assessment. Details of the imputation methods for missing data are given in CS Appendix Table 9.[a]	**No (but ERG conclude low risk of attrition bias in relation to this aspect of imbalances in missing data)** The FAS in this trial included all patients who underwent randomisation and received at least one dose of a study drug and had at least one assessment after baseline ("modified ITT" population). However, >97% of patients in the randomised population were included in the modified ITT population, suggesting risk of attrition bias would be low. The methods used to impute missing data appear appropriate.
[a] The methods used to account for missing data in the ITC have been transposed for the SCALE 1839 trial and STEP 1 trial in Appendix Table 9 compared to the descriptions in the ITC report. Source: partly reproduced from CS Appendix Table 9		

9.2 Results of other outcomes reported in the STEP 1 trial

This appendix reports the results of the following other outcomes measured in the STEP 1 trial: other weight loss outcomes, percentage of participants with a specified weight change from baseline, waist circumference change, incidence of type 2 diabetes (only reported at baseline and as a safety outcome), HbA_{1c} (%) change from baseline, and HRQoL

9.2.1 Weight loss

Body weight at baseline in kg was reported only for the FAS population. The mean (SD) baseline weight of participants was 104.5 (22.1) kg in the semaglutide 2.4 mg arm and 105.2 (21.5) kg in the diet and physical activity arm.

The change in body weight in kg from baseline to 68 weeks was only reported for the FAS population. The semaglutide 2.4 mg arm experienced a mean decrease of more than 15kg whilst the diet and physical activity arm experienced a mean decrease of less than 3 kg (Table 53). The difference between arms was statistically significant (95% CIs exclude zero).

Table 53 Weight change from baseline at 68 weeks

Estimand (Data source)	Semaglutide 2.4 mg		Diet and physical activity		Difference (95% CI)
	Mean change	N	Mean change	N	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Treatment policy (CSR 14.2.14)	-15.33 kg	1306	-2.61 kg	655	-12.71 kg (-13.68 to -11.74);p<0.0001
Trial product (CSR 14.2.22)	-17.36 kg	950	-2.70 kg	443	-14.66 (-15.58 to -13.74); p<0.0001
FAS: full analysis set					

The percentage of participants who achieved a weight decrease of ≥5%, ≥10%, ≥15% or ≥20% at 68 weeks relative to baseline was consistently higher in the semaglutide 2.4 mg arm than the diet and physical activity arm (Table 54). Odds ratios were statistically significant (95% CIs exclude 1.0) and were higher for the trial product estimand than for the treatment policy estimand (NB the trial policy estimand, which includes discontinuations and use of rescue medication, is likely to be more reflective of weight loss in clinical practice).

Table 54 Percentage of participants with specified weight change from baseline at 68 weeks (FAS population)

Estimand	Semaglutide 2.4 mg		Diet and physical activity		Odds ratio (95% CI)
	%	N	%	N	
Weight change ≥5%					
Treatment policy	86.4%	1212	31.5%	577	11.2 (8.9 to 14.2); p<0.001

Trial product	92.4%	1059	33.1%	499	37.0 (28.0 to 49.0)
Weight change ≥10%					
Treatment policy	69.1%	1212	12.0%	577	14.7 (11.1 to 19.4); p<0.001
Trial product	74.8%	1059	11.8%	499	30.0 (22.5 to 40.0)
Weight change ≥15%					
Treatment policy	50.5%	1212	4.9%	577	19.3 (12.9 to 28.8); p<0.001
Trial product	54.8%	1059	5.0%	499	31.8 (21.0 to 48.3)
Weight change ≥20%					
Treatment policy	32.0%	1212	1.7%	577	26.9 (14.2 to 51.0)
Trial product	34.8%	1059	2.0%	499	42.2 (20.8 to 85.6)
FAS: full analysis set Source: CS Figure 5 and trial publication					

9.2.2 BMI loss

The change in BMI from baseline to 68 weeks was reported only for the FAS population. The change was larger for the semaglutide 2.4 mg arm (a decrease of more than 5 kg/m^2) than for the diet and physical activity arm (a decrease of less than 1 kg/m^2), with the difference between arms statistically significant (95% CIs exclude zero) (Table 55).

Table 55 BMI change from baseline at 68 weeks

Estimand (Data source)	Semaglutide 2.4 mg		Diet and physical activity		Difference (95% CI)
	Mean change	N	Mean change	N	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Treatment policy (CSR 11.3.4.3)	-5.54 kg/m^2	1306	-0.92 kg/m^2	655	-4.61 (-4.96 to -4.27) (p NR)
Trial product (CSR 14.2.48)	-6.27 kg/m^2	1306	-0.95 kg/m^2	655	-5.33 (-5.65 to -5.00); p<0.0001
FAS: full analysis set; NR: not reported					

9.2.3 Waist circumference

Across the analyses conducted, mean waist circumference decreased from baseline to 68 weeks by 13.1 to 15.2 cm in the semaglutide 2.4 mg arm and by 4.1 cm to 6.1 cm in the diet and physical activity arm (Table 56). The difference between trial arms was statistically

significant for the FAS population (95% CIs exclude zero) but confidence intervals were not reported for the other analysis populations. The CS does not comment on the clinical significance of these changes in waist circumference, but the company explained in clarification response A20 that they are likely to be clinically meaningful. However, the ERG's clinical expert suggested that waist circumference is difficult to reliably measure in practice due to variations in waist shape and measurement errors, especially at higher BMIs. Waist circumference is not used in the company's economic analysis.

Table 56 Waist circumference change from baseline at 68 weeks

Estimand (Data source)	Semaglutide 2.4 mg		Diet and physical activity		Difference (95% CI)
	Mean (SD[a]) change	N	Mean (SD[a]) change	N	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Treatment policy (Appendix R.3)	-13.54 cm	1306	-4.13 cm	655	-9.42 (-10.30 to -8.53) (p NR)
Trial product (CSR 14.2.60)	-15.22 cm	1306	-4.48 cm	655	-10.75 (-11.6 to -9.88); p<0.0001
Target subgroup (BMI ≥30 plus ≥1 weight-related comorbidity) (post hoc analysis)					
Treatment policy (CS B.2.7.1)	-13.6 cm	974	-4.3 cm	496	-9.3 [b]
Trial product (Appendix E.2)	-15.22 (9.11) cm	974	-4.66 (9.28) cm	946	10.56 [b]
Liraglutide-eligible subgroup (BMI≥35 with non-diabetic hyperglycaemia and CVD risk) (post hoc analysis)					
Treatment policy (CS B.2.7.2)	-13.1 cm	273	-5.4 cm	148	-7.7 [b]
Trial product (Appendix E.2)	-14.69 (9.39) cm	273	-6.08 (9.52) cm	148	-8.61 [b]

FAS: full analysis set; NR: not reported
[a] SD reported for some analyses
[b] Not reported; raw difference calculated by reviewer

9.2.4 Incidence of type 2 diabetes

The incidence of type 2 diabetes is specified as an outcome in the NICE scope and decision problem but was reported only at baseline in the CS due to too few cases (ITC Report section 3.2). Although type 2 diabetes was an exclusion criterion in the STEP 1 trial, 18

patients in the semaglutide 2.4 mg arm (1.4%) and 17 in the diet and physical activity arm (2.6%) had type 2 diabetes at baseline (CS Table 8). According to the CSR, the incidence of diabetes in the STEP 1 trial safety analysis set was <0.1% (n=1 patient) in the semaglutide 2.4 mg arm and 0.9% (n=6 patients) in the diet and physical activity arm (CSR section 14.3.1.5).

9.2.5 Glycaemic status

Across the analyses conducted, mean % HbA_{1c} decreased from baseline to week 68 by 0.45 to 0.60 percentage points in the semaglutide 2.4 mg arm and by 0.1 to 0.2 percentage points in the diet and physical activity arm (Table 57). The difference between trial arms was statistically significant for the FAS population (95% CIs exclude zero) but confidence intervals were not reported for the other analysis populations.

The reductions in HbA1c in the semaglutide 2.4 mg arm were close to 0.50 %-points. The company's clarification response A20 states that according to clinical guidelines, a reduction of 0.5% (5.5 mmol/mol) is considered to be clinically significant (reference cited).

Table 57 HbA_{1c} (%) change from baseline at 68 weeks

Estimand (Data source)	Semaglutide 2.4 mg		Diet and physical activity		Difference (95% CI)
	Mean (SD[a]) change	N	Mean (SD[a]) change	N	
FAS (BMI≥30 or BMI≥27 plus ≥1 of hypertension, dyslipidaemia, OSA or CVD)					
Treatment policy (CSR 11.6.1)	-0.45 %-points	1306	-0.15 %-points	655	-0.29 (-0.32 to -0.26) (p NR)
Trial product (CSR 14.2.150)	-0.50 %-points	1306	-0.16 %-points	655	-0.34 (-0.37 to -0.31); p<0.0001
Target subgroup (BMI ≥30 plus ≥1 weight-related comorbidity) (post hoc analysis)					
Treatment policy (CS B.2.7.1)	-0.5 %-points	974	-0.1 %-points	496	-0.4 [b]
Trial product (Appendix E.2)	-0.52 (0.28) %-points	974	-0.17 (0.29) %-points	496	-0.35 [b]
Liraglutide-eligible subgroup (BMI≥35 with non-diabetic hyperglycaemia and CVD risk) (post hoc analysis)					
Treatment policy (CS B.2.7.2)	-0.5 %-points	273	-0.2 %-points	148	-0.3 [b]

Trial product (Appendix E.2)	-0.60 (0.28)	273	-0.16 (0.29)	148	-0.44 [b]
FAS: full analysis set; NR: not reported [a] SD reported for some analyses [b] Not reported; raw difference calculated by reviewer					

9.2.6 HRQoL outcomes

The HRQoL outcomes reported in STEP 1 are not used in the company's economic evaluation which instead draws upon HRQoL data from alternative sources considered more relevant to the longer time horizon of the economic model (see section 4.2.7). Only a brief summary of the STEP 1 HRQoL outcomes is therefore provided here.

Baseline HRQoL scores were comparable between the semaglutide 2.4 mg and diet and physical activity arms for each of the HRQoL scales assessed (trial publication).

In the FAS population the proportion of patients who achieved a clinically meaningful within-person improvement in HRQoL from baseline to week 68 was higher for the semaglutide 2.4 mg arm than the diet and physical activity arm when assessed using both the SF-36 physical functioning scale and the IWQOL-Lite-CT physical function scale (CS sections B.2.6.9.1 and B.2.6.9.2):

- SF-36 physical functioning (≥3.7 points): semaglutide 2.4 mg 40.0%, diet and physical activity 27.0%; OR=2.08 (95% CI 1.60 to 2.70)
- IWQOL-Lite-CT physical function (≥14.6 points: semaglutide 2.4 mg 51.2%, diet and physical activity 32.9%; OR=2.72 (95% CI 2.14 to 3.47)

However, the difference between semaglutide 2.4 mg and diet and physical activity was smaller for the improvement in the SF-36 Mental Component Summary (the company do not discuss the clinical significance of this) (trial publication).

9.3 Other ITC results

This appendix reports the results of outcomes from the ITC for the outcomes that are not used in the economic model. Note: none of the ITC results were used directly in the model.

9.3.1 BMI

BMI change was reported for STEP 1[5] and SCALE 1839,[19] but not included in the ITC analysis. We note, though, that percentage weight change and percentage BMI change are synonymous outcomes.

9.3.2 Waist circumference

The unadjusted analyses for both the treatment policy estimand and trial policy estimand indicate a statistically greater reduction waist circumference with semaglutide 2.4 mg than with liraglutide 3.0 mg except when the comparison was made at half a year (28 weeks in each trial) (Table 58). Adjusted analyses are only reported for the treatment policy estimand and these were also significantly in favour of semaglutide 2.4 mg except at the 28 weeks analysis. The treatment effect in unadjusted analyses was consistently larger for the trial product estimand than for the treatment policy estimand.

The CS does not discuss the clinical significance of the changes in waist circumference. Clarification response A20 states that after accounting for changes in BMI, reducing waist by 3 cm had a significant beneficial effect on the metabolic syndrome in women, and waist reductions of 5–10 cm in Caucasian women, across a range of baseline BMI 25–50 kg/m^2 or waist circumference 72–133 cm, may be used as guideline to encourage overweight women to achieve a realistic target with a high probability of health benefits (references cited). However, the company do not explicitly define a minimal clinically important change in waist circumference.

Table 58 ITC results: effect on waist circumference change from baseline

Analysis (STEP 1/SCALE 1839: week 52/56 unless stated)	**Relative treatment effect (95% CI), cm, semaglutide 2.4 mg vs liraglutide 3.0 mg**	
	Treatment policy estimand	**Trial product estimand**
Unadjusted	-3.59 (-5.56, -1.61), p=0.0004 [a,b]	-4.27 (-6.08, -2.45), p<0.0001 [b]
Population adjustment 1	-3.83 (-5.77, -1.88), p=0.0001 [b]	Not reported
Population adjustment 2	-3.75 (-5.72, -1.78), p=0.0002 [b]	Not reported
Unadjusted, pre-diabetes	-4.12 (-5.48, -2.76), p<0.0001 [b]	Not reported
Week 56/56, unadjusted	-3.57 (-5.54, -1.59), p=0.0004 [b]	Not reported
Week 68/56, unadjusted	-3.50 (-5.60, -1.40), p=0.0011 [b]	-4.47 (-6.39, -2.55), p<0.0001 [b]
Week 28/28, unadjusted	-0.59 (-2.15, 0.97), p=0.4586 [b]	-1.00 (-2.48, 0.48), p=0.1840 [b]

[a] From CS Table 12
[b] From ITC Report Table 10

9.3.3 HbA_{1c}

The unadjusted analyses for both the treatment policy estimand and trial product estimand indicate a statistically greater reduction in HbA_{1c} with semaglutide 2.4 mg than with liraglutide 3.0 mg (Table 59). Adjusted analyses are only reported for the treatment policy estimand and these were also significantly in favour of semaglutide 2.4 mg. The treatment effect in unadjusted analyses was similar for the trial product estimand than for the treatment policy estimand.

The reductions in HbA1c were relatively small (≤0.14 percentage point) but the CS does not discuss the clinical significance of these changes. In clarification response A20 the company state that in general, guidelines consider a difference of 0.5% (5.5 mmol/mol) to be clinically significant (reference cited).

Table 59 ITC results: effect on HbA_{1c} change from baseline

Analysis (STEP 1/SCALE 1839: week 52/56 unless stated)	**Relative treatment effect (95% CI), %-points, vs liraglutide**	
	Treatment policy estimand	**Trial product estimand**
Unadjusted	-0.13 (-0.20, -0.06), p=0.0002 [a,b]	-0.12 (-0.18, -0.06), p<0.0001 [b]
Population adjustment 1	-0.08 (-0.15, -0.01), p=0.0207 [b]	Not reported
Population adjustment 2	-0.08 (-0.15, -0.01), p=0.0324 [b]	Not reported
Unadjusted, pre-diabetes	-0.14 (-0.19, -0.09), p<0.0001 [b]	Not reported
Week 56/56, unadjusted	-0.13 (-0.19, -0.06), p=0.0002 [b]	Not reported
Week 68/56, unadjusted	-0.12 (-0.19, -0.05), p=0.0008 [b]	-0.13 (-0.18, -0.07), p<0.0001 [b]
Week 28/28, unadjusted	-0.06 (-0.12, -0.01), p=0.0293 [b]	-0.05 (-0.10, -0.00), p=0.0366 [b]

[a] From CS Table 12
[b] From ITC Report Table 9