NICE National Institute for Health and Care Excellence

Disease-specific reference case extension: management of overweight and obesity in adults

NICE process and methods
Published: 6 May 2026

www.nice.org.uk/process/pmg50

Contents

Introduction

This disease-specific reference case extension outlines standardised methods for health economic evaluations of interventions for managing overweight and obesity and reducing associated comorbidities in adults. Treatments that have multiple indications including for obesity are covered. This is to reflect that many people with obesity have 2 or more long-term health conditions.

This methods document should be used when undertaking cost-utility analyses to inform NICE guidance on overweight and obesity in adults (including guidelines, technology appraisal guidance and HealthTech evaluations). Elements of it may apply to other types of economic evaluations such as cost-comparison analyses.

Applicability of statements in this document should be decided at the scoping stage of guidance production. Adherence should be attempted while taking into account the specific decision problem under evaluation. Deviations should be highlighted at scoping stage and then explained and agreed through the committee's and NICE's quality assurance processes.

This document should also be read alongside the NICE manuals for:

- guidelines – particularly the section on the reference case for economic evaluations
- technology appraisals and highly specialised technologies – particularly the section on the reference case for economic evaluations
- HealthTech programme – particularly the section on economic evaluation.

For further details on developing and implementing disease-specific reference case extensions, including the basis for 'required' and 'recommended' statements and how these documents should be used, please refer to NICE's position statement on use of disease-specific reference models in economic evaluations and the glossary.

1 Population

1.1 Results need to be stratified by baseline presence or absence of type 2 diabetes mellitus (T2DM), baseline presence or absence of atherosclerotic cardiovascular disease (ASCVD) and whether the population is living with overweight or obesity at baseline (that is, reported separately and not combined – see table 1 for model stratification for presenting cost effectiveness results). For the definition of overweight and different classes of obesity, see table 2 on BMI categories by ethnic group.

Only the stratum or strata relevant to the intervention's proposed target population need be modelled.

When this population is more specific than any of the strata defined in table 1, only that specific population needs to be modelled. (required)

Table 1: Model stratification for presenting cost effectiveness results

Strata	Living with overweight or obesity at baseline?	Living with T2DM at baseline?	Living with ASCVD at baseline?
Stratum 1	Overweight	No	No
Stratum 2	Overweight	Yes	No
Stratum 3	Overweight	No	Yes
Stratum 4	Overweight	Yes	Yes
Stratum 5	Obesity	No	No
Stratum 6	Obesity	Yes	No
Stratum 7	Obesity	No	Yes
Stratum 8	Obesity	Yes	Yes

1.2 Further subgrouping should be undertaken to reflect differences in the baseline risk of events, treatment effect, and to capture existing thresholds for bariatric procedures and medications – see paragraphs 1.3 to 1.5. (recommended)

1.3 The population in strata 5 to 8 should be further subgrouped by body mass index (BMI) category (obesity class 1, obesity class 2, obesity class 3). The categorisation should be done by ethnicity when individual patient data about this is available as described in table 2 on BMI categories by ethnic group (source: NICE CKS: obesity diagnosis, identification and classification). (recommended)

Table 2: BMI categories by ethnic group

Category	BMI range: white (kg/m2)	BMI range: South Asian, Chinese, Middle Eastern, other Asian, Black African, and African-Caribbean (kg/m2)
Healthy weight	18.5 to 24.9	18.5 to 22.9
Overweight	25 to 29.9	23 to 27.4
Obesity class 1	30 to 34.9	27.5 to 32.4
Obesity class 2	35 to 39.9	32.5 to 37.4
Obesity class 3	40 or more	37.5 or more

1.4 For strata containing people without either T2DM or ASCVD, subgrouping by type or number of obesity-related comorbidities (such as dyslipidaemia, hypertension, metabolic dysfunction-associated steatotic liver disease (MASLD), obstructive sleep apnoea and non-diabetic hyperglycaemia [also referred to as 'pre-diabetes'] or metabolic syndrome) should be undertaken. (recommended)

1.5 For strata containing people with T2DM or ASCVD or T2DM and ASCVD, subgrouping by presence or absence of chronic kidney disease (CKD) or chronic heart failure (CHF) should be undertaken. (recommended)

1.6 Estimate baseline patient characteristics for each strata, such as age, sex, and weight, from real-world evidence representative of the NHS population in England. (required)

1.7 A sensitivity analysis should be conducted where baseline patient characteristics for each strata such as age, sex and weight, are estimated from trial data. (recommended)

For the rationale for these statements, see the rationale and supporting information section on population.

2 Intervention and comparators

2.1 Include relevant comparators in the model that are established practice in the NHS for managing weight, including those already recommended in NICE's guideline on overweight and obesity management. (required)

2.2 Behavioural interventions alone, with a focus on healthy diet and physical activity, are a relevant comparator. (required)

2.3 As a minimum, compare new medicines with medicine(s) established in NHS clinical practice for weight management (except where contraindicated). (required)

2.4 As a minimum, compare new bariatric procedures with bariatric procedures established in NHS clinical practice (except where contraindicated). (required)

2.5 As a minimum, compare new digital technologies with existing digital technologies established in NHS clinical practice for same use case(s). (required)

2.6 Bariatric procedures (with procedure type clearly defined and justified) should be comparators against other interventions targeted at the following subgroups:

- adults with a body mass index (BMI) of 30 kg/m² or more who have recent onset type 2 diabetes mellitus (T2DM), or
- adults with a BMI between 35 kg/m^2 and 39.9 kg/m^2 with a significant health condition that could be improved if they lost weight, or
- adults with a BMI of 40 kg/m² or more. (recommended)

  A lower BMI threshold (reduced by 2.5 kg/m^2) should be applied for people of South Asian, Chinese, other Asian, Middle Eastern, Black African or African–Caribbean background. (recommended)

2.7 If an intervention or comparator is indicated alongside a behavioural intervention, such as 'a healthy diet and physical activity', the duration of this behavioural intervention needs to reflect current NHS practice. Conduct sensitivity analyses

to reflect different durations and intensities. (required)

2.8 When comparing an intervention to behavioural interventions alone, explore different intensities of the behavioural intervention using sensitivity analyses. Include the lowest likely intensity of behavioural intervention of GP advice about diet and exercise once a year as part of the analyses to reflect the limited availability of services specifically for overweight and obesity. (required)

2.9 For the base case analysis, background treatments need to reflect established NHS practice relevant to the stratum, as detailed in the following NICE guidelines:

- overweight and obesity management
- type 2 diabetes mellitus in adults
- cardiovascular disease: risk assessment and reduction, including lipid modification. (required)

2.10 Conduct sensitivity analysis using the corresponding trial data when the background treatments in the trial does not reflect established NHS practice. (required)

For the rationale for these statements, see the rationale and supporting information section on intervention and comparators.

3 Model structure and health states

3.1 A state transition modelling approach, either as a cohort or individual patient-level simulation (IPS), is preferred (recommended). However, when alternative modelling approaches are used, these must be transparent and validated and capture all important differences in costs or outcomes between the intervention and comparator. (required)

3.2 Cycle length must not exceed 1 year. Time in each health state including associated costs and utilities will generally be half-cycle corrected. Front-loading of some high-cost items such as admissions or surgical procedures must be done if this is the only way to capture the acute resource use associated with the transition. (required)

3.3 Carefully select outcomes (health states and events) for each population group (stratum) guided by the following criteria (required):

- direct evidence of a treatment effect (from trials or real-world evidence), or strong evidence of an association of an improved outcome with weight loss and clinical plausibility that the intervention affects the outcome and
- evidence that the outcome has a meaningful impact on costs, quality of life or risk of other outcomes (for example, mortality)
- their inclusion within a high-quality core outcome set for that stratum, as these indicate outcomes that are important to patients and clinicians.

3.4 Capture health states in the model as follows:

- type 2 diabetes mellitus (T2DM) status: 'normoglycaemia' or 'non-diabetic hyperglycaemia' or 'T2DM' (required)
- atherosclerotic cardiovascular disease (ASCVD) acute events: myocardial infarction 'MI' or 'stroke' (ischaemic stroke or unspecified stroke) or 'stroke after MI' or 'MI after stroke' (required)
- ASCVD status: 'no ASCVD' or 'post-myocardial infarction' ('post-MI') or 'post-stroke' or 'post-stroke and post-MI' or 'ASCVD other' or 'post-MI and chronic

heart failure' ('post-MI and CHF') (required)

- chronic kidney disease (CKD) status: 'no CKD' or 'CKD stage G1A2 to G4' or 'CKD stage G5' (recommended)
- weight category: distinct health state ('healthy weight' or 'overweight' or 'obesity class 1' or 'obesity class 2' or 'obesity class 3') or using more granular weight or body mass index (BMI) trajectories (required)
- if applicable, line of weight management treatment – from line 1 to line 2 to line 3, etc (required)
- a 'dead' state (required).

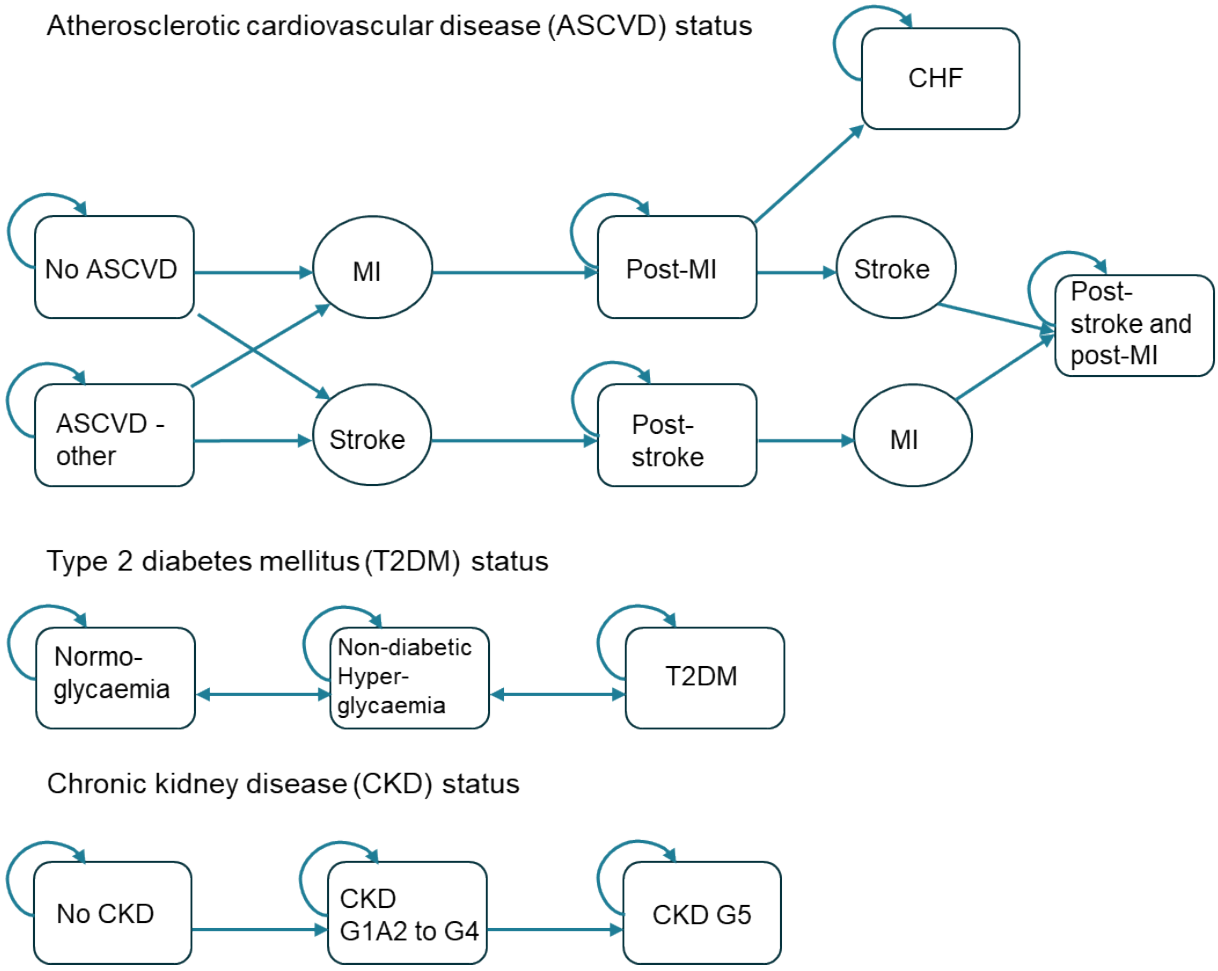
3.5 Strata containing people with T2DM at baseline should have additional health states containing diabetes-related core outcomes. (recommended)

3.6 Capture movement between health states as follows:

- remission and recurrence of non-diabetic hyperglycaemia (recommended)
- progression of T2DM status (from no diabetes to non-diabetic hyperglycaemia to diabetes) (recommended)
- progression of ASCVD status
  - from 'no ASCVD' or 'ASCVD other' to 'post-stroke' or 'post-MI' (recommended)
  - from 'post-stroke' or 'post-MI' to 'post-stroke and post-MI' (recommended)
  - from 'post-MI' to 'post-MI and CHF' (recommended)
- progression of CKD status (from 'no CKD' to 'CKD stage G1A2 to G4' to 'CKD stage G5') (recommended)
- change in weight category (required)
- all health states transition to the 'dead' state (required).

Key model transitions are illustrated in figure 1.

Figure 1: Key model state transitions



Definitions: CHF: chronic heart failure; MI: myocardial infarction.

3.7 Calculate costs and QALY losses for the following health events:

- obstructive sleep apnoea (recommended)
- osteoarthritis requiring knee replacement (recommended)
- osteoarthritis requiring hip replacement (recommended)
- bariatric procedure (unless this is an intervention or comparator in the model), with costs and utility decrements reflecting a weighted average of different procedures (recommended)
- treatment and procedure-related adverse effects (non-serious adverse

effects with an incidence of more than 5% for each intervention and any serious adverse effects) (required)

- discontinuation (required).

3.8 The following outcomes should be considered for inclusion as health states in a sensitivity analysis if there is direct evidence of a treatment effect:

- metabolic dysfunction-associated steatotic liver disease (MASLD) (recommended)
- cancer (breast, colorectal, kidney, liver and womb). (recommended)

For the rationale for these statements, see the rationale and supporting information section on model structure and health states.

4 Clinical parameters

Treatment effects

4.1 When available for all relevant comparators, use high-quality directly measured outcomes (for example, atherosclerotic cardiovascular disease [ASCVD] events or incidence of type 2 diabetes mellitus [T2DM]) to model relative treatment effects, including trends in treatment effects over time. (required)

4.2 Absolute effects should be modelled by multiplying the relative treatment effect by a baseline event rate. Baseline event rates should be estimated using real-world evidence when available. (recommended)

4.3 Baseline event rates should be specific to strata and health state and relevant to an NHS population in England, ideally also by age and sex subgroup. (recommended)

4.4 When directly measured health outcomes are unavailable or cannot be applied consistently across all relevant comparators, model the effectiveness of interventions through changes in weight and other risk factors, used as inputs to validated risk prediction tools or risk equations. (required).

4.5 When risk-factor-based modelling is used to predict changes in the incidence of ASCVD and T2DM, treatment-related changes in the following risk factors should be considered:

- weight or body mass index (BMI) (recommended)
- systolic blood pressure (recommended)
- HbA1c (recommended)
- cholesterol (high-density lipoprotein and low-density lipoprotein). (recommended)

4.6 Baseline estimates of risk factors (including weight) should be based on real

world evidence and be specific to stratum and health state and relevant to an NHS population in England. Treatment impact on risk factors should be based on evidence from trials. (recommended)

4.7 Justify the selection of validated risk prediction tools or risk equations to estimate modelled outcomes and events. (required)

4.8 Risk prediction tools applicable to the stratum or subgroup and to the NHS context in England should be used. (recommended)

4.9 When using multi-year risk estimates (such as 10-year risks) convert them into annual probabilities precisely to prevent overestimation. (required)

4.10 When data for high-quality directly measured health outcomes is available for some (but not all) comparators, use it to validate, and when necessary, calibrate, outcomes predicted using risk equations or risk prediction tools. (required)

4.11 Even when direct evidence is used to measure health outcomes, weight (baseline and treatment effect) should be modelled over time. This is for the adjustment of utilities based on weight. (recommended)

For the rationale for these statements, see the rationale and supporting information section on treatment effects.

Treatment effects in the longer term

4.12 Predictions of treatment effects beyond trial follow-up need to be biologically and clinically plausible, for example, by taking into account data on the natural history of weight change, and long-term adherence to treatment. (required)

4.13 Long-term weight trajectory assumptions across behavioural intervention, pharmacological intervention (allowing variation by mechanism of action), minimally invasive and surgical bariatric procedures should be differentiated to reflect their distinct clinical benefits. This may be informed by real-world longitudinal data. (recommended)

4.14 When extrapolation beyond the data observation period is needed, incorporate the following scenario analyses to explore uncertainty in weight trajectory over time while on treatment, including:

- applying the same natural history weight gain to both interventions and comparators (required)
- assuming no natural history weight gain in interventions and comparators (required)
- exploring partial waning of treatment effect relative to the initial magnitude of treatment benefit on weight loss over time. (required)

4.15 When weight is not used to predict health outcomes, incorporate scenarios around intervention treatment effect waning in a sensitivity analysis for these outcomes. (required)

For the rationale for these statements, see the rationale and supporting information section on treatment effects in the longer term.

Treatment discontinuation

4.16 Include treatment duration or discontinuation and its effect on resource use in the model for interventions and comparators, including behavioural intervention only comparators. (required)

4.17 Beyond trial follow-up, reduce relative treatment effect to reflect discontinuation or dose reduction, as treatment effects should apply only to people continuing with the (full-dose) treatment. (required)

4.18 The heterogeneity across people in post-treatment weight regain should be incorporated to capture the long-term health outcomes adequately. (recommended)

4.19 When medicines have been discontinued, do not include NHS resource use associated with concomitant behavioural interventions unless there is evidence

that behavioural interventions without medicine have continued in clinical practice in England. (required)

4.20 Treatment discontinuation should be modelled over time as follows:

- rate of discontinuation over time should be specific to the population stratum and treatment and relevant to an NHS population in England (recommended)
- baseline discontinuation rates should be sourced from real-world evidence and relative differences in discontinuation rates between interventions from trial data. (recommended)

4.21 When evaluating interventions that have a waning of the treatment effect, threshold analyses should be included to identify optimal treatment durations. (recommended)

4.22 Account for weight regain and subsequent impact on health states and events after treatment discontinuation using evidence from long-term studies when available:

- when long-term studies for the intervention are not available, use evidence on weight regain for other treatments with a similar mode of action and efficacy or clinical expert opinion (required)
- use rate of weight gain over time that is specific to the population stratum and treatment and relevant to an NHS population in England (required)
- include scenario analyses informed by clinical opinion to test the impact of different weight regain rates or duration of regain, for example, return to initial baseline weight after a specific number of years of discontinuation (required)
- include a scenario where weight returns to a higher level than the underlying natural trajectory of weight gain on no treatment (required)
- include a scenario where weight returns to pre-treatment weight in a shorter timeframe than in the base case analysis. (required)

For the rationale for these statements, see the rationale and supporting information section on treatment discontinuation.

Mortality

4.23 Use a consistent mortality modelling approach within the model structure; ideally mortality rates should be specific to both the health state and the body mass index (BMI) category. If specific data is not available, then the model needs to either apply BMI-adjusted all-cause mortality ratios or condition-specific mortality ratios (not both) to avoid the risk of double counting. (required)

4.24 When possible, mortality rates (cardiovascular mortality and non-cardiovascular mortality) should be specific to the population stratum and health state and relevant to an NHS population in England. (recommended)

4.25 BMI-adjusted mortality should be applied to non-CVD mortality only, with CVD mortality modelled separately through ASCVD health states. Alternatively, if using BMI-adjusted all-cause mortality, CVD death should not be modelled as a separate outcome. When BMI-adjusted mortality is used, granularity in BMI categories is preferred: age-sex-specific hazard ratios reflect the non-linear and age-dependent relationship between BMI category and mortality risk. (recommended)

4.26 When all-cause mortality evidence is available from trials use it to validate and, if necessary, calibrate the overall mortality effect predicted in the model. (required)

4.27 Sensitivity analyses should be used to explore the impact of length of time of living with overweight or obesity on mortality. (recommended)

For the rationale for these statements, see the rationale and supporting information section on mortality.

5 Measuring and valuing health effects

5.1 Measure utility using EQ-5D based on age, body mass index (BMI) and sex and ensure these are adjusted over time based on age and BMI. Apply utility decrements to specific events and comorbidities. (required)

5.2 The best source for utilities is a single dataset that controls for weight, comorbidity and other variables such as age and sex. If that is not available then, to avoid double-counting, ensure that the source study for weight-related utilities is controlled for comorbidity or the source study for comorbidity-related utilities is controlled for weight. (required)

5.3 Mean EQ-5D scores for each intervention in the clinical trials should be used to calibrate the mean quality-of-life treatment effect in the short-term (at trial follow-up), ensuring not to double count health-related quality-of-life improvements from the reduced incidence of progression or adverse effects. (recommended)

5.4 Capture utility decrements for treatment-related adverse effects and complications from bariatric procedures in the model. Do this by weighting the incidence rate of the adverse effect by the duration of the event. For example, include gastrointestinal adverse effects such as nausea, diarrhoea and constipation for medicines, and capture recovery time and complications for bariatric procedures. (required)

5.5 When there is weight regain after an intervention, a sensitivity analysis should be conducted to model a greater decline in utility than the increase in utility associated with the initial weight loss. (recommended)

For the rationale for these statements, see the rationale and supporting information section on measuring and valuing health effects.

6 Cost and healthcare resource use

6.1 Include incremental costs associated with assessment for eligibility for any intervention, such as suitability for bariatric surgery procedures. (required)

6.2 Include costs of complications, including the costs of re-operation and re-intervention for bariatric procedures. (required)

6.3 Include the following implementation costs:

- supply chain costs associated with medicines, if applicable (required)
- capital investment costs associated with bariatric procedures, if applicable. (required)

6.4 For behavioural interventions, whether as a comparator or used concomitantly with other interventions, ensure resource use reflects current NHS practice in the base case. In a sensitivity analysis, base this on resource use associated with the relevant trials. (required)

6.5 Health state and event costs, including costs associated with treatment or procedure-related adverse effects, should align with clinical practice and, when appropriate, be specific to the population stratum and be sourced from the NHS in England. (recommended)

6.6 Resource use for all health states should capture use of mental healthcare services, because this is important for all morbidities. (recommended)

6.7 The best source for background costs is a single dataset that controls for weight, comorbidity and other variables such as age and sex. If that is not available then, to avoid double-counting, ensure that either the source study for weight-related costs is controlled for comorbidities or the source study for comorbidity-related costs is controlled for weight. (required)

6.8 Resource use for modelled comorbidities should consider duration of comorbidity, with different costs for managing newly-diagnosed comorbidities

versus those for managing established comorbidities. (recommended)

6.9 For people with type 2 diabetes mellitus (T2DM), management and complication costs from cohort studies should be used. For example, the UK Prospective Diabetes Study (UKPDS), which include the cost of consultations, visits, admissions and procedures associated with diabetes-related complications. (recommended)

For the rationale for these statements, see the rationale and supporting information section on cost and healthcare resource use.

7 Equality and other considerations

7.1 If there is clear evidence of a significant burden of health inequalities, a distributional cost-effectiveness analysis (DCEA) could be undertaken with results presented, as supporting evidence, by deprivation quintile groups. If a DCEA is undertaken, this should be by stratum and variation in take-up and adherence to treatment across deprivation group should be captured as well as variation in prevalence of disease. This analysis could be particularly helpful to encourage uptake of interventions that may support tackling health inequalities. (recommended)

7.2 Do not include cost-effectiveness results by these deprivation quintile groups as part of the base-case analysis or present them as non-reference case scenarios. (required)

7.3 Some benefits and risks might be difficult to fully capture quantitatively in the modelling. Therefore, consider qualitatively the impact of the intervention on the following:

- health inequalities related to protected characteristics, for example, South Asian ethnicity or disability such as severe mental illness, autism or learning disability, because incidence of obesity is higher in these groups (recommended)
- access to other treatments, for example, organ transplants (recommended)
- treatment-related outcomes, for example, fertility (recommended)
- the health of unborn children (recommended)
- in the case of bariatric procedures, how recovery time impacts carers, workers and those from deprived areas more than others. (recommended)

For the rationale for these statements, see the rationale and supporting information section on equality and other considerations.

Rationale and supporting information

Population

Sections 1.1 to 1.7

Modelling needs to reflect that people with specific comorbidities will have different levels of risk of health events and capacity to benefit from treatments for overweight and obesity and will be on different care pathways.

Type 2 diabetes mellitus (T2DM) and atherosclerotic cardiovascular disease (ASCVD) were identified as key comorbidities for people living with overweight and obesity. Therefore, it was agreed that results should be stratified according to whether people have these conditions. This also reflects how treatments are recommended in NICE's guideline on type 2 diabetes in adults and NICE's technology appraisal guidance on tirzepatide for treating type 2 diabetes and semaglutide for reducing the risk of major adverse cardiovascular events in people with cardiovascular disease and overweight or obesity.

Further subgrouping of the obesity strata by body mass index (BMI) category was considered appropriate by clinical experts because:

- the risk of developing comorbidities is linked to BMI category
- it allows for any differences in treatment effect by BMI category to be identified and so
- allows for recommendations to be developed for NICE guidance that capture this.

Clinical experts and patient advocates debated using BMI compared to waist circumference for stratifying the population of interest. They highlighted the importance of waist circumference measurement and its use in waist-to-height ratio calculations, particularly for people with a BMI of less than 35 kg/m^2 as recommended in NICE's guideline on overweight and obesity management, but noted that it is unlikely to be useful for modelling purposes as currently there are no risk equations that use this measure to estimate risk of developing associated comorbidities.

Stratifying obesity in adults into 3 groups as outlined in the section on classifying obesity in adults in NICE's guideline on overweight and obesity management is desirable. However,

that would have increased the total number of strata from 8 to 16. To keep it manageable, only stratification between living with overweight and living with obesity is 'required'. Subgrouping the obesity strata further, although desirable, might not be essential for all cost-utility analyses and so this has been rated as 'recommended'.

Ethnicity-specific BMI thresholds were included to reflect evidence reported in NICE's guideline on overweight and obesity management that certain ethnic populations (for example, South Asian, Chinese, Middle Eastern, Other Asian, Black African and African-Caribbean) have a higher cardiometabolic risk at lower BMI levels than white populations. When individual patient data is available, it was agreed that subgrouping should be done by ethnicity, taking into account the ethnic-specific BMI thresholds. The healthy weight category does not need be included as a baseline weight category, but it is relevant for later modelling.

Stratification by both BMI category and comorbidity allows the impact of interventions on cardiovascular events and T2DM to be captured.

The number and type of comorbidities in the population are important as they influence baseline risk of morbidity and mortality. Treatment may result in a greater reduction in risk for certain subgroups. Sensitivity analyses that explore the cost-effectiveness of interventions in different population subgroups defined by both type and number of comorbidities will help identify the populations that will gain the greatest benefit.

Populations without T2DM or ASCVD should be grouped using other comorbidities such as dyslipidaemia, hypertension, metabolic dysfunction-associated steatotic liver disease (MASLD), obstructive sleep apnoea and non-diabetic hyperglycaemia (in line with NICE's technology appraisal guidance on semaglutide and tirzepatide for managing overweight and obesity), and metabolic syndrome. However, it was considered these might have less power to predict health outcomes within the 6 population strata that have T2DM or ASCVD or T2DM and ASCVD, where the cost, quality of life and prognosis will already be poorer, and so subgrouping using these additional comorbidities was not recommended.

Chronic kidney disease (CKD) and chronic heart failure (CHF) have been included as subgroups for the population strata containing people with T2DM or ASCVD or T2DM and ASCVD. This was to align with modelling undertaken for the medicines update of NICE's guideline on type 2 diabetes in adults and to reflect the increased prevalence of these conditions in these populations (Dawson et al. 2023, Koye et al. 2018, Lee et al. 2024, Panchal et al. 2024).

It is acceptable to only include the strata or stratum that reflect the eligible target population for the intervention of interest. For example, if an intervention is targeting people with a BMI of 40 kg/m^2 or more and ASCVD, then only this stratum should be presented.

There are various causes of genomic and syndromic obesity, each of which has unique features and treatment pathways that will not be adequately covered by this reference case extension. These include Prader-Willi Syndrome, Bardet-Biedl Syndrome and Alström Syndrome. A broad range of population groups may be affected by obesity. However, as outlined in NICE's position statement, the primary purpose of disease-specific reference case extensions is to promote consistency in health economic modelling across NICE assessments. To achieve this, reference case extensions are expected to focus on common diseases and conditions where multiple pieces of NICE guidance typically exist, and where standardisation can therefore provide the greatest value.

Populations in trials of interventions for obesity may not reflect the full eligible population in England. Therefore, baseline patient characteristics, such as age, sex and weight, for each strata should be estimated from real-world evidence representative of the NHS population in England. A sensitivity analysis should be undertaken when patient characteristics are estimated from trial populations, to assess generalisability of the trial results to a real-world population whose condition is managed in the NHS in England.

Intervention and comparators

Sections 2.1 to 2.10

To ensure consistency and relevance across cost-utility analyses, relevant comparators that are established practice in the NHS for managing weight, including those already recommended by NICE, should be included. These comparators should be relevant to the stratum being evaluated. These comparators represent the spectrum of treatment intensity and allow for meaningful comparisons of cost-effectiveness.

The inclusion of bariatric procedures as a comparator reflects NICE's guideline on overweight and obesity management and acknowledges the clinical and cost-effectiveness of surgical interventions for people with obesity. This has been included as a 'recommended' rather than 'required' statement as the strength of NICE's recommendations on assessment for bariatric procedures in its guidance vary, depending on the body mass index (BMI) of the population. There are several bariatric procedures

available in the NHS. These vary in terms of their invasiveness, reversibility, effectiveness, complications and costs. As well as surgical procedures, there are also minimally invasive endoscopic bariatric interventions (NICE's HealthTech guidance on endoscopic sleeve gastroplasty for obesity). The comparator procedure should be clearly defined in the model to ensure appropriate effectiveness data and costs are applied. Justification for the choice of bariatric procedure should be provided and should reflect established NHS practice. Currently the most common bariatric procedures in NHS practice are sleeve gastrectomy and Roux-en-Y pass (National Obesity Audit's bariatric surgical procedure dashboard). The model should cover the provision of other interventions while people are waiting to receive bariatric procedures.

In line with clinical practice and NICE guidelines, if an intervention is indicated alongside a behavioural intervention, then this should be captured in the model. These interventions may be delivered in person or digitally. However, the intensity and structure of behavioural interventions can vary significantly, making it challenging to achieve consistency between models. To address this, different intensities of behavioural interventions should be explored in sensitivity analyses, with each behavioural intervention clearly described and costed. This should be done separately for the concomitant and standalone behavioural interventions.

When comparing an intervention to behavioural interventions alone, sensitivity analyses should include the cost of the minimal behavioural intervention of GP advice once a year to reflect the limited availability of services for overweight and obesity in the NHS. Such a sensitivity analysis only explores the lower cost of a minimal intervention of annual GP advice and does not explore any potential treatment effect. This is a conservative analysis that aims to provide an upper estimate of the incremental cost effectiveness ratio (ICER).

For strata that include people with type 2 diabetes mellitus (T2DM) or atherosclerotic cardiovascular disease (ASCVD), standard background treatment for those conditions should be included and costed, ensuring that models reflect realistic treatment pathways and avoid underestimating costs or overestimating incremental benefits. In the base case these background treatments should reflect those used in the NHS.

Initial treatment for T2DM included modified-release metformin and an sodium glucose transport 2 inhibitor (SGLT-2 inhibitor), except where contraindicated (see NICE's guideline on type 2 diabetes in adults). For people with ASCVD, background treatment should include lipid-lowering treatment and other relevant treatments. However, background treatment in trials may not reflect established clinical practice in the NHS, and sensitivity

analysis should reflect background treatments used in the corresponding trial evidence to ensure they are aligned with the effectiveness data.

Model structure and health states

Sections 3.1 to 3.8

A state transition modelling approach, either cohort-based or individual patient-level simulation (IPS), was considered suitable by health economic modellers.

State transition models are well-suited to chronic disease modelling and have been widely used in obesity and related comorbidities (see appendix B). They offer relative transparency to more complex modelling approaches, ease of use and can capture disease progression over time. State transition models can be used for a wide range of intervention types for managing obesity, such as medicines, surgery and behavioural interventions, and allow for isolation of specific benefits when needed to support transparent decision-making.

NICE considers state transition models to be appropriate for modelling this condition. However, discrete event simulations offer some advantages and so are potentially acceptable if the model is transparently presented and can be easily interrogated.

IPS models allow for tracking of individual patient histories, which is particularly important in obesity because events such as treatment discontinuation, bariatric procedures or cardiovascular events can occur at varying times and influence future risks. IPS models can better capture dependencies between events, which are difficult to represent in memoryless cohort models. If a cohort model is used, tunnel states should be added to address these limitations.

Obesity is associated with a wide range of health conditions, some causally linked and some potentially contributing to obesity itself, such as polycystic ovary syndrome. Certain comorbidities, such as type 2 diabetes mellitus (T2DM) and atherosclerotic cardiovascular disease (ASCVD), especially stroke and myocardial infarction (MI), can significantly alter treatment pathways and long-term outcomes. Therefore, economic models should include a carefully selected set of comorbidities and health events that reflect the burden of obesity. The inclusion of comorbidities and health events should be guided by the following criteria: the need for a strong, evidence-based association with obesity, that they have a substantial impact on costs or quality of life, and their inclusion in a high-

quality core outcome set (COS). The association should be identified by measuring the change in the outcome relative to a reduction in weight or body mass index (BMI) – see, for example, Khunti et al. 2023.

When there are several core outcomes sets (COSs), the quality of the COS development process can be assessed using the Core Outcome Set-STAndards for Development (COS-STAD) criteria (Kirkham et al. 2017). A recent report in the COMET database (Gorst et al. 2025) concluded that the COS developed by Coulman et al. (2016) was methodologically sound, meeting 10 of the 12 COS-STAD criteria, and with a good number of participants with lived experience throughout the process. Therefore, all of the outcomes in that COS are included in this reference case extension.

A minimum set of health states to be included in economic model have been provided. Only events or outcomes that have been directly measured or use validated risk equations (see rationale and supporting information section on treatment effects for details on using validated risk equations) should be included. These should reflect the most common and impactful obesity-related comorbidities.

Chronic kidney disease (CKD) was included to ensure consistency with the modelling undertaken for the medicines update of NICE's guideline on type 2 diabetes in adults. The link between obesity and CKD is well established, and there is some evidence that incretin agonists may improve renal outcomes via both direct effects (including glucose lowering in those with diabetes and other tissue effects) and indirect effects (via weight loss). However, the magnitude and mechanism of effect remain uncertain, and ongoing trials such as REMODEL (Cherney et al. 2025) or SURMOUNT-MMO (Lam et al. 2025) may clarify this.

Transitions in the model should reflect realistic disease trajectories to ensure clinical credibility and avoid overestimating intervention benefits. They are essential because they reflect the natural history of obesity and its complications. The included transitions are supported by robust evidence and are key drivers of long-term costs and health outcomes.

Other events, while not needing to be captured as health states in the economic model, have significant cost and utility implications and so should be included as health events in the model. These were identified by reviewing previous obesity health economic models (as outlined in appendix B) and after discussions with clinical experts and patient advocates. The following health events should be included for the reasons given:

- obstructive sleep apnoea – because it is prevalent in people with obesity and contributes to reduced quality of life and increased healthcare utilisation
- knee and hip replacement – because these are costly treatments for osteoarthritis, a common weight-related condition
- bariatric procedures – because of its substantial impact on weight, comorbidities and long-term costs
- all serious adverse effects and non-serious adverse effects that occur in more than 5% in each intervention – based on clinical trial reporting standards and partly reflecting the inclusion criteria in NICE's technology appraisal guidance on tirzepatide for managing overweight and obesity.

The inclusion of these events, as well as discontinuation of medication, should help ensure that models capture the full burden of obesity and the real-world consequences of treatment. They are all 'recommended' for inclusion, except for treatment and procedure-related adverse effects, which is 'required' because, unlike the other events, it is relevant for all assessments.

Some health events are recommended for inclusion in sensitivity analyses if there is emerging or partial evidence of a treatment effect. These include:

- Metabolic dysfunction-associated steatotic liver disease (MASLD): Included in the model for NICE's technology appraisal guidance on tirzepatide for managing overweight and obesity due to trial data but not included in NICE's technology appraisal guidance on liraglutide or semaglutide for managing overweight and obesity due to lack of demonstrated impact. Inclusion should be evidence-driven to avoid speculative modelling. NICE is currently developing technology appraisal guidance for treating people with MASLD and associated liver fibrosis with resmetirom or semaglutide.
- Cancer: While obesity is a known risk factor for several cancers, there is limited evidence demonstrating a reduced risk of cancer following weight loss. In NICE's technology appraisal guidance on liraglutide for managing overweight and obesity, when cancer was included in the economic model, the incremental cost effectiveness ratio (ICER) increased slightly, because the cost of cancer treatment in the extra months of life was greater than the benefits of a modest reduction in the incidence of cancer. The reverse was true for a scenario with a subgroup with a high risk of cancer at baseline. Either way, the inclusion of cancer is not recommended unless there is

evidence of a reduction in the incidence of cancer with weight loss.

These events were not included in the minimum health state set because of insufficient direct evidence of treatment effect, uncertainty in causal pathways and potential for modelling complexity or bias. This flexible approach ensures that models remain evidence-based, while allowing opportunity to incorporate new data as it becomes available.

Several additional obesity-related health outcomes such as fertility, depression, chronic lower back pain and inflammatory conditions were discussed at workshops involving clinical experts, patient advocates, health economic modellers, industry stakeholders and commissioners. They were not included in the minimum health states or events because there was insufficient direct evidence of treatment effect and concerns with the potential for modelling complexity and double counting of impact on quality of life. It was agreed that these additional health outcomes could be captured qualitatively. However, if new direct evidence of treatment effect emerges to warrant their inclusion, then they can be considered for inclusion within future versions of this reference case extension.

It was noted that for populations with MASLD at baseline, consideration of additional health states may be necessary. Some guidance is available specifically for metabolic dysfunction-associated steatohepatitis (MASH) – see NICE's health technology assessment innovation laboratory report on evaluating MASH treatments.

Procedure-related adverse events that are important to patients include (Coulman et al. 2016):

- technical complications (including leaks, fistulas, strictures, ulcers, intraoperative organ injury and internal hernia)
- reoperation or reintervention
- problems with dysphagia and regurgitation.

Clinical parameters

Treatment effects

Sections 4.1 to 4.11

When available for all relevant comparators, directly measured health outcomes should be used to model relative treatment effectiveness. This should be high-quality direct evidence from randomised control trials (RCTs), but when RCTs evidence is unavailable or inappropriate, real-world evidence may be used.

Potential bias and confounding factors in real-world evidence should be appropriately addressed. Some studies, including both RCTs and real-world observational studies, report low numbers of outcome events. Although such studies may provide direct evidence of treatment effect, the fact that they only report a low number of outcome events leads to imprecise effect estimates, often reflected in wide confidence intervals. This lack of precision increases uncertainty around the size of the treatment effect and, in some cases, even its direction. As a result, this evidence cannot be considered high quality or sufficiently reliable for decision making, regardless of study design.

Relative effects can be estimated from RCTs, or network meta-analyses of RCTs when there are more than 2 interventions being compared. Treatment effect may not be constant over time; it may wane over time due to reduced rates of adherence to medication over time, or due to the mechanism of the intervention, and so analyses should incorporate changes in treatment effect over time where possible.

Baseline event rates (for example, for populations that have not received a bariatric procedure or been exposed to incretin agonist use) should be specific to strata and health state, as there is demonstratable heterogeneity in atherosclerotic cardiovascular disease (ASCVD) and non-ASCVD population groups, especially in terms of the risk of further ASCVD events and life expectancy, which have a subsequent impact on costs and quality of life. These rates should be estimated from real-world evidence that is relevant to an NHS population in England, to ensure that the predictions of the cost-effectiveness analyses are generalisable. Absolute event rates for each intervention should be estimated by multiplying relative treatment effects to this baseline event rate. While it may not be possible to estimate subgroup-specific treatment effects for each relevant subgroup, model developers should clearly state the assumptions applied when subgroup-specific estimates are unavailable. Testing for heterogeneity of effect between strata is recommended.

However, when such directly measured health outcomes are unavailable or cannot be applied consistently across interventions, effectiveness should be modelled in terms of changes in weight and other risk factors (for example, blood pressure, HbA1c levels and cholesterol levels) that are linked to the health outcomes and health events included in the

model structure. This includes situations when directly measured outcomes are available for some interventions but not others, and therefore cannot be applied consistently across comparators. Trials may only report limited intermediate health outcomes and lack long-term data, which is often the case for behavioural and digital interventions. In such situations, it is necessary and appropriate to estimate treatment effects using validated risk equations or prediction tools. A summary of risk equations and risk prediction tools used in previous obesity models is reported in appendix B.

In this context, risk-factor-based modelling provides a pragmatic means of estimating long-term health outcomes when direct evidence is limited, rather than replacing directly observed health outcomes. These risk factors are aligned with the cardiometabolic risk outcomes highlighted in the International Consortium of Health Outcomes Measurement's patient-centred outcome measures for adults living with obesity. This approach is pragmatic, widely used in health economic modelling, and allows for the estimation of long-term health and cost outcomes based on short-term or intermediate-term clinical data.

When available, directly measured health outcomes should be used to validate and, when necessary, calibrate predicted health outcomes from risk equations. Direct evidence of intervention effect on comorbidity progression (such as ASCVD and type 2 diabetes mellitus [T2DM]) or health events is particularly important for interventions that may have effects beyond weight loss, such as medicines that influence metabolic or cardiovascular outcomes through additional mechanisms (Sattar et al. 2025).

Active registries, for example the National Bariatric Surgery Registry and National Obesity Audit, or NHS data accessed via secure data environments may be appropriate sources of real-world evidence depending on the specific data needs.

Emerging evidence suggests that the length of time living with overweight and obesity has a significant impact on future health outcomes, particularly for conditions such as ASCVD, T2DM, and osteoarthritis (Krüger et al. 2025, Zeng et al. 2023). People with these conditions may carry a residual risk from having them even after substantial weight loss, due to irreversible physiological damage accumulated over time. For example, long-term obesity can accelerate atherosclerosis and joint degeneration, which are not fully reversed by subsequent weight reduction. This is especially relevant for mechanical outcomes like osteoarthritis, where cumulative joint stress from prolonged high body mass index (BMI) may lead to persistent damage. Therefore, ideally risk prediction tools should be based on changes in risk factor over time (longitudinal data) rather than on data on risk factors

measured at a single point in time for a population (cross-sectional data). Current risk prediction tools often rely on cross-sectional BMI measurements and may not adequately capture the long-term impact of living with overweight and obesity.

When data is available, the impact of length of time living with overweight and obesity on future health outcomes should be explored in sensitivity analyses, for example, by distinguishing between people who have recently started living with obesity and those that have lived with obesity for much longer.

Treatment effects in the longer term

Sections 4.12 to 4.15

When direct evidence is used to model health outcomes, predictions of treatment effects beyond trial follow-up should be biologically and clinically plausible (for example, if levels of treatment adherence are expected to reduce over time, it will be less realistic to assume ongoing treatment effects based on the assumption of treatments being fully adhered to). When selecting methods for extrapolation, for example, if survival analysis is used, guidance in the NICE DSU Technical Support Document 14 should be followed.

Risk equations are a pragmatic and widely used approach in obesity modelling. However, they are not designed for causal inference and often do not capture the dynamic nature of treatment effects, such as initial weight loss followed by weight regain. They may also underestimate clinical benefits, particularly for interventions with indirect or multi-system effects (for example, cardiovascular or metabolic improvements independent of weight loss).

Multiple scenario and sensitivity analyses should be conducted to explore the impact of different assumptions on model outcomes, as there is considerable uncertainty around long-term effects including weight trajectories, especially beyond trial follow-up periods. These should explore both conservative and optimistic assumptions to give an indication of a range of cost-effective estimates.

Many interventions show an initial weight reduction followed by partial weight regain (Ahmed 2024, Wu et al. 2025). Ignoring this trajectory can misrepresent long-term effectiveness. Models should capture treatment effects over time such as the rate of weight change relative to the initial weight over time and weight regain patterns, to better reflect the fact that weight loss is rarely sustained at the same level. The treatment effect

waning should also be measured by modelling the percentage change in bodyweight from baseline and applying the hazard ratio afterwards.

Treatment discontinuation

Sections 4.16 to 4.22

Models should incorporate treatment duration or discontinuation for all treatment comparators, including 'behavioural intervention only' comparators. The baseline rates should be taken from real-world evidence and relative differences between interventions from trial data, when available, and reflect their impact on weight and future outcomes. If high-quality real world evidence is available, the baseline discontinuation should refer to the rate observed in any comparator arm (not limited to behavioural intervention only) and trial-derived treatment effect should be applied to that estimate.

Discontinuation occurs for several reasons, such as patient self-directed discontinuation or stopping of treatment because of adverse effects (common in medicines for weight management), treatment inefficacy or achievement of healthy weight. The reason for discontinuation may be associated with the duration of treatment benefit, and the likelihood of weight regain, but it is unlikely that it will be possible to incorporate this data into economic modelling, given that discontinuation rates are not often recorded in databases or reported in the literature in this way.

Commissioners highlighted that when medicines are discontinued, the provision of concomitant behavioural interventions is also stopped. The costs in the model should reflect this unless evidence of these behavioural interventions continuing without medicine is available.

There was a strong emphasis from clinical experts and commissioners on the value of real-world evidence to validate assumptions about treatment duration, discontinuation and long-term effectiveness. This is because clinical trials often differ from routine practice in terms of population characteristics, adherence and duration of treatment.

NICE's technology appraisal guidance on semaglutide for managing overweight and obesity found the medicine to be less cost-effective after 2 years because of its waning effect on weight. Given this, threshold analyses are recommended for new treatments to find the point at which treatment is no longer cost-effective. Treatment effects in the longer term should reflect changes in dose as well as discontinuation and general

treatment waning effects.

The purpose of capturing discontinuation rates in the model is in part to capture resource use. However, beyond the trial follow-up it should also be used to modify the treatment effect. If most people are likely to have stopped using the medicine, then the treatment effects would not be expected to continue beyond the duration of the trial follow-up. The impact of treatment discontinuation on changes to health outcomes such as glycaemic control should therefore be accounted for. A systematic review (West et al. 2026) found that weight regain after treatment with weight management medications was faster than after treatment with a behavioural intervention. Therefore, the impact of treatment discontinuation on weight should be explored comprehensively in the model.

Threshold analyses should only be included in economic analyses to identify optimal treatment durations for interventions under evaluation. This is because it might not be cost-effective to continue treatment when there is waning of treatment effect over time. Within technology appraisals or HealthTech evaluations, this should be applied to new medicines or technologies that are the subject of the appraisal (but not for established treatments that are not being evaluated), or when updating guidance if it is in the remit of the evaluation.

Mortality

Sections 4.23 to 4.27

Mortality should be modelled in a way that is methodologically robust and reflective of the underlying epidemiological evidence. A consistent approach should be adopted, whereby models either apply body mass index-adjusted (BMI-adjusted) all-cause mortality ratios or condition-specific mortality ratios. Combining both risk ratios may lead to double-counting and should be avoided unless clearly justified.

To reduce the potential risk of under- or over-estimating mortality effects, model-predicted mortality should be validated against all-cause mortality observed in clinical trials. Where necessary, using observed trial data to calibrate the aggregate mortality predicted by the model.

Models should recognise that the length of time people have lived with overweight or obesity may increase the risk of mortality even after weight reduction. Assuming full reversal of BMI-related mortality risk may overestimate treatment benefits. Therefore,

sensitivity analysis is recommended.

Measuring and valuing health effects

Sections 5.1 to 5.5

Models submitted in NICE's technology appraisal guidance on liraglutide, semaglutide and tirzepatide for managing overweight and obesity used baseline utility based on age, body mass index (BMI) and sex, with utility decrements applied to specific events.

A study by Luah et al. in 2024 estimated the association between BMI and EQ-5D-5L among the general population in England using data from 2017 and 2018 health surveys. It distinguished utility values by sex and BMI level. It also derived the coefficients for comorbidities, including for diabetes, heart and circulatory disease, respiratory disease, musculoskeletal disease, cancer and mental health disorders. This may be a suitable study to inform the baseline utility values and utility decrements associated with comorbidities in a consistent manner.

EQ-5D data from patients in relevant clinical trials should be used to calibrate the quality-of-life treatment effect in the short-term (at trial follow-up), as direct evidence of quality-of-life improvement is best when it is precise and of high quality.

Treatment-related adverse effects and complications from bariatric procedures should be estimated by weighting by incidence and effect duration, using an evidence-based estimate of duration.

The most common treatment-related adverse effects reported for incretin agonists were gastrointestinal. Clinical experts at the workshop noted this was a key reason for dose titration of medicines. Such adverse effects occur early and are unlikely to be ongoing because those experiencing problems after minimising the dose will stop taking the medicine.

In the case of bariatric procedures, the model should reflect the fact that quality of life will be reduced while patients are recovering from their procedure. In the case of minimally-invasive procedures, this time is likely to be less than for surgical procedures but will depend on complication rates.

Utility decrements for treatment-related adverse effects were applied additively in base

cases, and a multiplicative approach has been explored in scenario analyses in previously published technology appraisal guidance (see appendix B for further details). Health economic modellers noted that the most appropriate approach to applying health state utility estimates depends on whether or not the health effects are seen to be independent.

If the analysis by Luah et al. (2024) is used to estimate health state utility values, an additive approach is most appropriate as the utility decrement values in the study were estimated with a linear regression.

If other sources are identified for health state utilities and the utility decrements associated with type 2 diabetes mellitus (T2DM), obesity and any other comorbidities show no significant interactions, then an additive approach also seems reasonable, otherwise a multiplicative approach as discussed in DSU Technical Support Document 12 should be used. This document also outlines methods for adjusting utility values when calibration is required, such as for quality-of-life effect in the short-term (at trial follow-up), and for avoiding double counting improvements from the reduced incidence of T2DM progression or adverse effects.

Obesity is associated with mental illness. This can worsen when lost weight is regained, and so strategies that produce sustainable weight loss will be most effective at improving rates of mental illness associated with obesity. This was an important health outcome highlighted by clinical experts and patient advocates. Impact on mental health should generally be captured by the EQ-5D if it is used in trials or if weight-related EQ-5D estimates are used in the model, although it may not capture elements around the social stigma of obesity. For strategies where there is weight regain, the subsequent loss of utility will not be fully captured by simple weight-related utilities. Therefore, we recommend sensitivity analysis for weight regain.

Cost and healthcare resource use

Sections 6.1 to 6.9

Interventions for managing weight require assessment for eligibility and can be associated with complications, for example, adverse effects for medicines or re-operation or re-intervention for procedures. The costs associated with both assessment and managing complications should therefore be captured in the model.

For medicines, supply chain costs may include dispensing, refrigeration, waste disposal

and delivery costs. For bariatric procedures, capital investment costs, including infrastructure costs, should be captured (see NICE technology appraisal and highly specialised technologies guidance: the manual and NICE HealthTech programme manual). Costing should consider device reuse parameters, organisational impact (such as training requirements or investments in infrastructure) and learning-curve effects, when relevant, consistent with Drummond et al. (2018).

As noted by the committee for NICE's technology appraisal guidance on tirzepatide for managing overweight and obesity, there is currently uncertainty about the resource use needed for behavioural interventions (related to encouraging a healthy diet and physical activity), both as comparators and concomitant treatments with other interventions. NHS England's interim commissioning guidance on implementation of NICE's technology appraisal guidance on tirzepatide states that wraparound care for tirzepatide should incorporate appropriate nutritional and dietetic advice, physical activity guidance and behavioural change components, over a minimum timeframe of 9 months from the point of prescribing. The importance of exploring the resource use for these behavioural interventions in sensitivity analyses was highlighted during the workshops. If real-world evidence is available about implementing tirzepatide (for example, data collected as part of NHS England's interim commissioning guidance on implementation of NICE's technology appraisal guidance on tirzepatide) in terms of the resource use required for concomitant behavioural interventions, then this should be used.

For the cost of comorbidities and acute events, adjustments should be made to ensure that these costs are not double counted. For example, the cost of treating obesity should be subtracted from the cost of treating type 2 diabetes mellitus (T2DM) or atherosclerotic cardiovascular disease (ASCVD), where relevant. When multiple comorbidities develop sequentially, ideally costs should be measured incrementally from the same data set. However, as a minimum, acute event care costs should be greater than costs associated with chronic states. Clinical experts emphasised the importance of capturing the impact of weight management on mental health and therefore there is a requirement to include the cost of mental healthcare services, as the impact on mental health itself should already be captured by the EQ-5D.

The cost of treating comorbidities often increases over time as more additional comorbidities develop. With T2DM, for example, average resource use is likely to underestimate costs for advanced disease and overestimate costs for incident cases. Workshop experts and the committee for NICE's technology appraisal guidance on tirzepatide for managing overweight and obesity advise that attempts should be made to

distinguish the costs of early and later disease.

For people with T2DM, management and complication costs from cohort studies are recommended. For example, the committee for NICE's technology appraisal guidance on tirzepatide for managing overweight and obesity preferred using UK Prospective Diabetes Study (UKDS) costs, which include inpatient and non-inpatient healthcare costs as a function of T2DM-related complications, age and sex, rather than the company's initial approach of using diabetes-related NHS reference costs and assuming 1 hospital attendance or stay per patient per year. The cost of medicines is not included in the UKPDS costing and therefore would need to be costed separately.

Equality and other considerations

Sections 7.1 to 7.3

Obesity is more prevalent in the most socioeconomically deprived quintiles of the population (see NICE's health inequalities briefing report for its guideline on overweight and obesity management and its health inequalities report for its medicines update of its guideline on type 2 diabetes in adults).

Treatments for obesity can potentially help reduce the differences in health outcomes between different socio-economic groups if uptake is particularly supported in those groups.

Distributional cost-effectiveness analysis (DCEA) can show the potential impact of a weight management intervention on health inequalities, which could support the case for recommending the intervention (see the section on distributional cost-effectiveness analysis methods in NICE's health technology evaluations manual). However, the number of adults with obesity in England is very large and so committees should be aware that the opportunity cost associated with recommending an intervention with a slightly higher incremental cost effectiveness ratio (ICER) could be substantial. DCEA can also be used to indicate when additional recommendations or targeted implementation support might improve take-up of treatments in deprived areas and strengthen the equity impact.

There is some evidence that take-up of, and adherence to, weight management treatments might be lower in more deprived groups. This should be captured in DCEAs – see NICE's inequality analysis of health outcomes of different diets in achieving and maintaining weight loss for NICE's guideline on overweight and obesity management

(January 2025).

The methods for DCEA for aspects of health inequality other than socioeconomic deprivation are less well developed. Therefore, these should be considered outside of the model.

For pragmatic reasons, models inevitably do not cover every effect. A few elements that are unlikely to be captured in economic models were identified as important by clinical experts, patient advocates, health economic modellers, industry stakeholders and commissioners. These have been highlighted for qualitative consideration outside of the model.

Health inequalities for people living with disability and for specific ethnic groups have been highlighted based on the NICE's health inequalities briefing report for its guideline on overweight and obesity management and the advice of workshop experts.

Weight management is a key consideration in provision of certain treatments. Two treatments where this is the case were identified in the expert workshops: fertility treatment (Boddeti et al. 2025, Pandey 2010) and transplant surgery (Ghanem et al. 2024). Others include dialysis, atrial fibrillation ablation and surgery more generally. It should be noted that if weight loss enables other treatments to go ahead, then it also incurs the associated costs of these treatments, which can be considerable, for dialysis, in particular.

Real-world data analysis to inform modelling

Table 3 shows data analysis opportunities that can help with model inputs that meet this reference case extension or supporting information about model design.

Any real-world data analysis should follow the principles set out in the NICE real-world evidence framework. When estimating baseline rates, these might be for a "behavioural intervention" arm that focuses on healthy diet and physical activity, and for a population that is not having bariatric procedures or incretin agonists. Data should be relevant to an NHS population in England; data from NHS populations across the UK are considered generalisable and can be used.

Table 3: Useful real-world data analyses

Parameters or analysis	Details	Examples of suitable data sets
Baseline population characteristics	Extract age-sex, weight and comorbidity distributions by stratum in the prevalent population	GP medical records (for example, Clinical Practice Research Datalink or OpenSafely) linked to hospital episode statistics (HES)
Choosing health outcomes and comorbidities	Test association between weight loss or gain and health outcome changes	GP medical records linked to HES and general registry data
Baseline risk of health events and outcomes	Estimate baseline rates of all-cause mortality, acute clinical events and incident comorbidities by stratum (defined in table 1), and preferably by age-sex group and obesity subgroup	GP medical records linked to HES and Office for National Statistics (ONS) Civil Registrations of Death

Parameters or analysis	Details	Examples of suitable data sets
Risk equations for health events and outcomes	Conduct multivariate analysis that includes age, sex and risk factors (weight or body mass index [BMI], systolic blood pressure, HbA1C and cholesterol level), by stratum (defined in table 1), and preferably by age-sex group and obesity subgroup	GP medical records linked to HES
Baseline weight change over time	Estimate weight change over time, by stratum (defined in table 1) and preferably by age-sex group and obesity subgroup	GP medical records
Weight change after discontinuation	Estimate BMI change over time after discontinuation of incretin agonists or after bariatric procedures, by stratum (defined in table 1), and preferably by age-sex group, obesity subgroup, and reason for discontinuation	GP medical records
Procedure-related complications	Estimate the proportion of people with complications for each procedure, and the number of complications for each procedure, as people may experience more than one complication per procedure	National registries (for example, National Bariatric Surgery Registry) or GP medical records linked to HES
Health state utilities	Multivariate analyses to determine EQ-5D, with BMI, age, sex and various comorbidities or health events as covariates	Registries or surveys (for example, Health Survey for England)
Health state costs	Multivariate analyses to determine healthcare cost, with BMI, age, sex and various comorbidities or health events as covariates	GP medical records linked to HES, combined with unit costs from standard UK sources (for example, National Cost Collection).
Resource use of behavioural interventions	Type, frequency and duration of healthcare professional contacts in addition to any other resource use (for example subscriptions)	GP medical records or surveys

Further research on data inputs is required to populate models for genetic or syndromic obesity.

Methodology

To inform the development of this disease-specific reference case extension for obesity, the following activities were conducted:

- a pragmatic review of recent health economic modelling approaches used in the evaluation of obesity interventions, drawing on NICE guidance, publications from other health technology assessment agencies and academic literature (see appendix B)
- consideration of the epidemiological data on the prevalence of obesity-related comorbidities to support decisions on which conditions should be included in economic models and how they should be represented.

In addition, NICE convened a series of virtual engagement workshops to explore key methodological issues and preferred modelling approaches. Each workshop targeted a specific stakeholder group to ensure a broad range of perspectives:

- first workshop engaged clinical experts and patient organisations to understand clinical pathways and patient priorities
- second workshop focused on health economic modellers, discussing technical modelling challenges and potential solutions
- third workshop brought together industry stakeholders to explore the implications of modelling choices on innovation and market access
- fourth workshop was held with commissioners to gain their perspective on the reference case extension.

Insights from these workshops, along with feedback from a wider stakeholder consultation, informed the statements in this report. A detailed overview of the workshops is provided in appendix A.

Glossary

Terms used in this document are defined in the NICE glossary and below.

Atherosclerotic cardiovascular disease (ASCVD)

Current or previous history of peripheral arterial disease, coronary artery disease such as myocardial infarction and unstable angina, cerebrovascular disease such as transient ischaemic attack and ischaemic stroke or unspecified stroke. Definition obtained from NICE's guideline on type 2 diabetes in adults. The health state 'ASCVD other' refers to current or previous history of peripheral arterial disease or transient ischaemic attack.

Bariatric procedures

A term used to encompass surgical (invasive or minimally invasive) and non-surgical procedures (minimally invasive).

Baseline

1). The starting time-point of the model, usually the decision-point with respect to choice of treatment strategy. The opposite of the 'time horizon', baseline characteristics include the age and comorbidities of patients at the outset.

2). 'Baseline risks' are the event rates in 1 arm of the model. The event rates in another model arm might be determined by multiplying the baseline risk with a relative treatment effect.

Cohort model

A type of health economic model that follows a group of people with shared characteristics through defined health states over time, using average transition probabilities to estimate cost effectiveness of interventions.

Direct evidence

When direct evidence of an effect is required, this should be evidence from randomised controlled trials (RCTs) for the specific treatment. When RCT evidence is unavailable or inappropriate, direct evidence from a real-world setting may be used, subject to appropriate assessment of bias and confounding factors.

Individual patient-level simulation

A modelling approach that simulates outcomes for individual patients rather than a whole cohort. Each simulated patient has unique characteristics and pathways through health states, allowing variability in risks, events and treatment effects to be captured.

Partial waning

Used when treatment effectiveness declines over time to a lower level.

Recommended statements

Where there is some uncertainty, an approach is 'recommended' within this reference case extension. These statements should be adhered to, as appropriate, and when possible. When deviations occur, these should be discussed with the committee and NICE quality assurance staff.

Required statements

Where there is strong evidence and it is considered a best approach, this reference case extension makes it clear that a specific element is 'required'. These statements must be adhered to unless there is a strong justification for not doing so, such as new evidence that contradicts the guidance in this document. When deviations occur, these should be discussed and agreed with the committee and NICE quality assurance staff.

Risk equations

Risk equations are mathematical formulas used to estimate the probability of a health event based on patient characteristics. These equations are often derived from regression

models using clinical trial or cohort data and are commonly used in economic models to simulate disease progression or treatment impact.

Risk prediction tools

These use 1 or more risk equations and other types of approaches, such as algorithms and scores, to provide an estimate of a person's risk of a health event.

State transition model

A modelling approach in which a population or individual moves between defined health states over certain time periods. State transition models can be implemented as cohort models or individual-level models.

Stratum

A stratum is a group of people that is pre-defined at baseline based on specific characteristics. Stratification is used to ensure that each group is adequately represented, and each stratum is mutually exclusive. The reference case extension defines 8 strata for use in NICE economic evaluations of weight management. Cost-effectiveness results should not be pooled across different strata.

Subgroup

A subgroup is a set of participants, used to explore differences in health outcomes or effects across specific characteristics (for example, sex or comorbidity status). Subgroups can overlap and are often used in subgroup analyses.

Threshold analysis

An analysis used to identify the value of a key input at which an intervention becomes cost-effective.

Tunnel state

A temporary health state in a state transition (Markov) model used to capture time-

dependent events or short-term effects that cannot be represented by standard Markov states.

List of appendices (separate documents)

Appendix A: Workshop notes

Appendix B: Health economic literature review

Appendix C: Scope

Appendix D: Equality and health inequality assessment (EHIA)

Appendix E: Acknowledgements

References

Ahmed I.A.B. (2024) A comprehensive review on weight gain following discontinuation of glucagon-like peptide-1 receptor agonists for obesity. Journal of Obesity, 2024(1), 8056440.

Ara R., Wailoo A.J. (2011) NICE DSU Technical Support Document 12: The use of health state utility values in decision models.

Boddeti S., Noor K., Ansar M., Abera M., Suresh S.B., Malireddi A., Khan S. (2025) Impact of obesity on in vitro fertilization (IVF) outcomes: a systematic review. Cureus, 17(8), e90695.

Cherney D.Z., Belmar N., Bjornstad P., Chacko M.M., Gunnarsson T.P., Hodgin J.B., Kretzler M., Pruijm M., Schytz P.A., Tuttle K.R. (2025) Rationale, design and baseline characteristics of REMODEL, a mechanism-of-action trial with semaglutide in people with type 2 diabetes and chronic kidney disease. Nephrology Dialysis Transplantation.

Coulman K.D., Hopkins J., Brookes S.T., Chalmers K.A., Main B.G., Owen-Smith A., Hughes R., Blazeby J.M. (2016) A core outcome set for the benefits and adverse events of bariatric and metabolic surgery: the BARIACT study. PLoS Medicine, 13(11), e1002187.

Dawson L.P., Carrington M.J., Haregu T., Nanayakkara S., Jennings G., Dart A., Stub D., Kaye D. (2023) Differences in predictors of incident heart failure according to atherosclerotic cardiovascular disease status. ESC Heart Failure, 10(6), 3398–3409.

Drummond M.F., Tarricone R., Torbica A. (2018) Economic evaluation of medical devices.

Ghanem O.M., Pita A., Nazzal M., Johnson S., Diwan T., Obeid N.R., Croome K.P., Lim R., Quintini C., Whitson B.A., Burt H.A. (2024) Obesity, organ failure, and transplantation: a review of the role of metabolic and bariatric surgery in transplant candidates and recipients. Surgical Endoscopy, 38(8), 4138–4151.

Gorst S., Williamson P.R., Kirkham J.J. (2025) COMET Initiative study record: Core outcome set development.

Khunti K., Schnecke V., Haase C.L., Harder-Lauridsen N.M., Rathor N., Sommer K., Morgen C.S. (2023) Weight change and risk of obesity-related complications: a retrospective

population-based cohort study of a UK primary care database. Diabetes, Obesity and Metabolism, 25(9), 2669–2679.

Kirkham J.J., Gorst S., Altman D.G., Blazeby J., Clarke M., Devane D., Gargon E., Williamson P.R. (2017) Core Outcome Set-STAndards for Development: The COS-STAD recommendations. PLoS Medicine, 14(11), e1002447.

Koye D.N., Magliano D.J., Nelson R.G., Pavkov M.E. (2018) The global epidemiology of diabetes and kidney disease. Advances in Chronic Kidney Disease, 25(2), 121–132.

Krüger A., van Dijk K.W., van Heemst D., Noordam R. (2025) Long-term body mass index trajectories and the risk of type 2 diabetes mellitus and atherosclerotic cardiovascular disease using healthcare data from UK Biobank participants. Atherosclerosis, 403, 119135.

Lam C.S.P., Rodriguez A., Aminian A., Ferrannini E., Heerspink H.J.L., Jastreboff A.M., Laffin L.J., Pandey A., Ray K.K., Ridker P.M., Sanyal A.J., Yki-Jarvinen H., Mason D., Strzelecki M., Bartee A.K., Cui C., Hurt K., Linetzky B., Bunck M.C., Nissen S.E. (2025) Tirzepatide for reduction of morbidity and mortality in adults with obesity: rationale and design of the SURMOUNT-MMO trial. Obesity.

Latimer, N. (2011) NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.

Lee H.S., Lim H.I., Moon T.J., Lee S.Y., Lee J.H. (2024) Trajectories of atherosclerotic cardiovascular disease risk scores as a predictor for incident chronic kidney disease. BMC Nephrology, 25(1), 141.

Luah X.W., Holst-Hansen T., Lübker C. (2024) The association between body mass index and health-related quality of life in the 2017 and 2018 Health Survey for England data: a cross-sectional observational analysis. Diabetes, Obesity & Metabolism, 26(6), 2318–2328.

Panchal K., Lawson C., Chandramouli C., Lam C., Khunti K., Zaccardi F. (2024) Diabetes and risk of heart failure in people with and without cardiovascular disease: systematic review and meta-analysis. Diabetes Research and Clinical Practice, 207, 111054.

Pandey S., Pandey S., Maheshwari A., Bhattacharya S. (2010) The impact of female obesity on the outcome of fertility treatment. Journal of Human Reproductive Sciences,

3(2), 62–67.

Sattar N., Lee M.M.Y. (2025) Estimating direct tissue effects versus weight loss effects of incretin-based drugs for obesity on various chronic conditions. The Lancet Diabetes & Endocrinology, 13(4), 347–354.

West S., Scragg J., Aveyard P., Oke J.L., Willis L., Haffner S.J.P., Knight H., Wang D., Morrow S., Heath L., Jebb S.A., Koutoukidis D.A. (2026) Weight regain after cessation of medication for weight management: systematic review and meta-analysis. BMJ, 392, e085304.

Wu H., Yang W., Guo T., Cai X., Ji L. (2025) Trajectory of body weight after drug discontinuation in the treatment of anti-obesity medications. BMC Medicine, 23(1), 398.

Zeng M., Chen S., Fan T., Yang H., Cao P., Wang Z., Zhang Y., Zhang Y., Hunter D., Yang Q., Ding C. (2023) Associations of childhood-to-adulthood body size trajectories and genetic susceptibility with the risks of osteoarthritis: a population-based cohort study of UK Biobank data. The Lancet Global Health, 11, S2.

ISBN: 978-1-4731-9488-5