National Institute for Health and Care Excellence

Report version (Draft for consultation)

Medicines associated with dependence or withdrawal symptoms: safe prescribing and withdrawal management for adults

Cost-utility analysis: Group CBT alongside tapering off and tapering off alone for people taking benzodiazepines

NICE guideline <number>

Economic analysis report

October 2021

Draft for Consultation

This analysis was developed by the National Guideline Centre

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and, where appropriate, their carer or guardian.

Local commissioners and providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

Contents

1

1 Introduction

Benzodiazepines are psychoactive drugs that, together with Z-drugs, are indicated and widely used for the short-term relief of severe or disabling anxiety associated with insomnia or short-term psychosomatic, organic, or psychotic illness.[8]

Although the British National Formulary (BNF) states that benzodiazepines should not be prescribed for more than 4 weeks, anecdotal data suggest that most people use benzodiazepines for much longer. A recent UK study[13] found that approximately a quarter of a million people in England are continuously taking benzodiazepines and that some of them are willing to discontinue the medication if adequately helped.

Benzodiazepine use is associated with a range of adverse effects, the most common being over-sedation and impaired cognitive function[53] which may lead to long-term cognitive problems in the elderly.[29] In addition, some people report increased aggressive behaviour which has been associated with self-harm or suicide[9] or harm to others. Due to their sedating properties, benzodiazepines have been also associated with road-traffic accidents and fall injuries in the elderly.[5, 15]

Discontinuing benzodiazepines is not an easy task as clinical evidence has clearly shown that severe withdrawal complications can be experienced even in people receiving lower doses.[53] Currently, few dedicated NHS centres provide support to people who want help with withdrawal[13] implying that many who wish to stop benzodiazepines are unable to find the necessary support. This is likely to have long-term implications on NHS resources and public health given the large and robust literature on the harms of long-term use of benzodiazepines.

There are several treatment options available for those willing to discontinue benzodiazepines. The evidence collected during the guideline clinical review suggests that a tapered withdrawal programme supported by a general practitioner provides a clinically important benefit in the number of people who successfully discontinue their benzodiazepine and that such a programme is even more effective if offered alongside cognitive behavioural therapy (CBT). However, the cost effectiveness of CBT and tapered withdrawal programmes compared to usual care remains uncertain in England.

A cost-effectiveness analysis[48] conducted alongside one of the trials found tapering off (TO) + CBT and tapering off alone (TOA) to reduce benzodiazepine costs compared to usual care but was deemed insufficient to inform a recommendation. Even though the overall clinical review found CBT effective, this trial found no effectiveness of CBT and, consequently, the results of its economic analysis were considered not reliable. In addition, as the benefits of discontinuing benzodiazepines occur in the long-term, the committee thought that an economic evaluation should look at the lifetime benefits of discontinuing benzodiazepines and that any recommendation should not rely on a within-trial analysis only.

The cost per person of a group CBT session and of a tapered withdrawal programme were estimated to be around £150 and £242, respectively. A study[13] suggests that in England around 120,000 people may be willing to accept help to withdrawal from benzodiazepines, therefore a recommendation on CBT and tapered withdrawal will likely have a significant resource impact on the NHS. Hence, the need to develop a model to assess whether tapering withdrawal with or without CBT is cost-effective compared to usual care in England appears to be strongly justified.

2 Methods

2.1 Model overview

A cost-utility analysis was undertaken where lifetime quality-adjusted life years (QALYs) and costs from a current UK NHS and personal social services perspective were considered. The analysis followed the standard assumptions of the NICE reference case for interventions with health outcomes in an NHS setting including discounting at 3.5% for costs and health effects[38] An incremental analysis was undertaken.

2.1.1 Comparators

The following comparators were included in the analysis:

1. Group cognitive behavioural therapy plus tapering off (CBT+TO)
2. Tapering off alone (TOA)
3. Usual care (UC)

Usual care was defined as the continuation of standard therapy with benzodiazepines with no attempt to reduce or discontinue it.

2.1.2 Population

The population of the analysis was people who had been using benzodiazepine for over 3 months and are unable to quit their usage by themselves. The average age and proportion of females were taken from the trials included in the clinical review although in the sensitivity analysis two different age groups were tested (see initial cohort settings in table 3 of section 2.3.2). Initially, conducting a subgroup analysis on people with polysubstance or substance misuse disorder, was discussed, as evidence suggested that they may have a higher all-cause mortality, but the committee ultimately decided not to include this subgroup as many polysubstance users may be under illicit or non-prescribed medicines and, therefore, they are out of scope of this guideline.

2.2 Approach to modelling

The difference in the cessation rate with CBT+TO, TOA and usual care was based on the evidence identified in the guideline clinical review. A Markov model was developed to estimate long-term health outcomes (QALYs) and costs for each comparator. Benefits of the two interventions aiming at discontinuing benzodiazepines were captured as reduced benzodiazepine consumption and diminution of long-term adverse effects of benzodiazepines. The model was run for 50 cycles representing 50 years of life as some of the consequences of benzodiazepine use, for example cognitive impairment, occur in older age. Time spent in each health states and the number of adverse events occurring over the lifetime of people were calculated to determine costs and QALYs associated with each strategy. The comparison between the results of each strategy allowed us to identify the most cost-effective strategy.

Details of the Markov model structure are described in section 2.2.1. To account for uncertainty, a probabilistic analysis was undertaken (see section 2.2.3 for further details).

2.2.1 Model structure

The structure of the Markov model is illustrated in figure 1. People enter the model either in the “abstinent” or “On benzodiazepine” states in a proportion determined by the effectiveness

of the strategy considered. People in the "abstinent" state have a positive probability of relapsing during the first two cycles and transiting to the "on benzodiazepine" state if they start taking the drug again. Relapse is allowed only during the first two cycles representing the first 2 years following the intervention, as the committee's clinical experience and published literature[14] suggest that most of the relapses occur during the first 2 years and that being abstinent at 21 months is strongly associated with being abstinent at further follow-up (up to 10 years). However, a sensitivity analysis was included with relapsing occurring for 5 years.

People both in the "abstinent" and "on benzodiazepines" states are at risk of experiencing adverse events with a higher probability in the benzodiazepine state reflecting the role of benzodiazepines in causing these events. After a non-systematic literature review and discussion with the committee, the following adverse events of benzodiazepines were incorporated:

1. Long-term cognitive impairment (dementia)
2. Hip fracture
3. Fall injuries
4. Suicide
5. Road traffic accident

Dementia and hip fracture were modelled as long-term Markov states whereas the other events were not modelled explicitly as states but as events occurring at each cycle and depending on the population at risk.

Dementia is associated with a specific cost occurring at the offsetting of the condition necessary for the correct diagnosis and with long-term health care costs and losses of utility lasting for the rest of the life of the patients. Therefore, dementia was modelled with 2 distinct Markov states: a tunnel state reflecting the diagnostic phase of the disease and the state "dementia long-term" capturing costs and consequences of the diseases occurring in the long-term. People in the tunnel state spend one cycle only in such state before transiting to "dementia long-term". As there is no definitive cure for dementia, it is assumed that once a person enters the "dementia long-term" state, they cannot leave it unless they die.

Hip fracture is associated with an acute phase when the person receives the treatment for the fracture and is at a very high risk of dying, and with a chronic phase lasting at least up to 10 years after the offsetting of the condition characterized by NHS costs, increased mortality relative to the general population and lower quality of life. As such, hip fracture was modelled in two states to capture both the acute and chronic phases. "Hip fracture acute" is a tunnel state where people transit before moving to the chronic "post-hip fracture" state and is associated with the highest mortality and cost for the NHS. "Post-hip fracture" is considered a permanent state, so the model does not allow people to transit from this state, unless they die. Although it is unclear whether hip fracture affects health permanently, published literature suggests that mortality is affected for a period of at least 10 years[23] and that a hip fracture, perhaps through an inflammatory or immunologic effect, may trigger or accelerate frailty in patients with few comorbid conditions at baseline, causing long-lasting effect perhaps for the rest of life[31]. A further discussion on the potential issues arising from using a permanent health state for hip fracture is presented in section 2.3.5.2 and section 4.2.

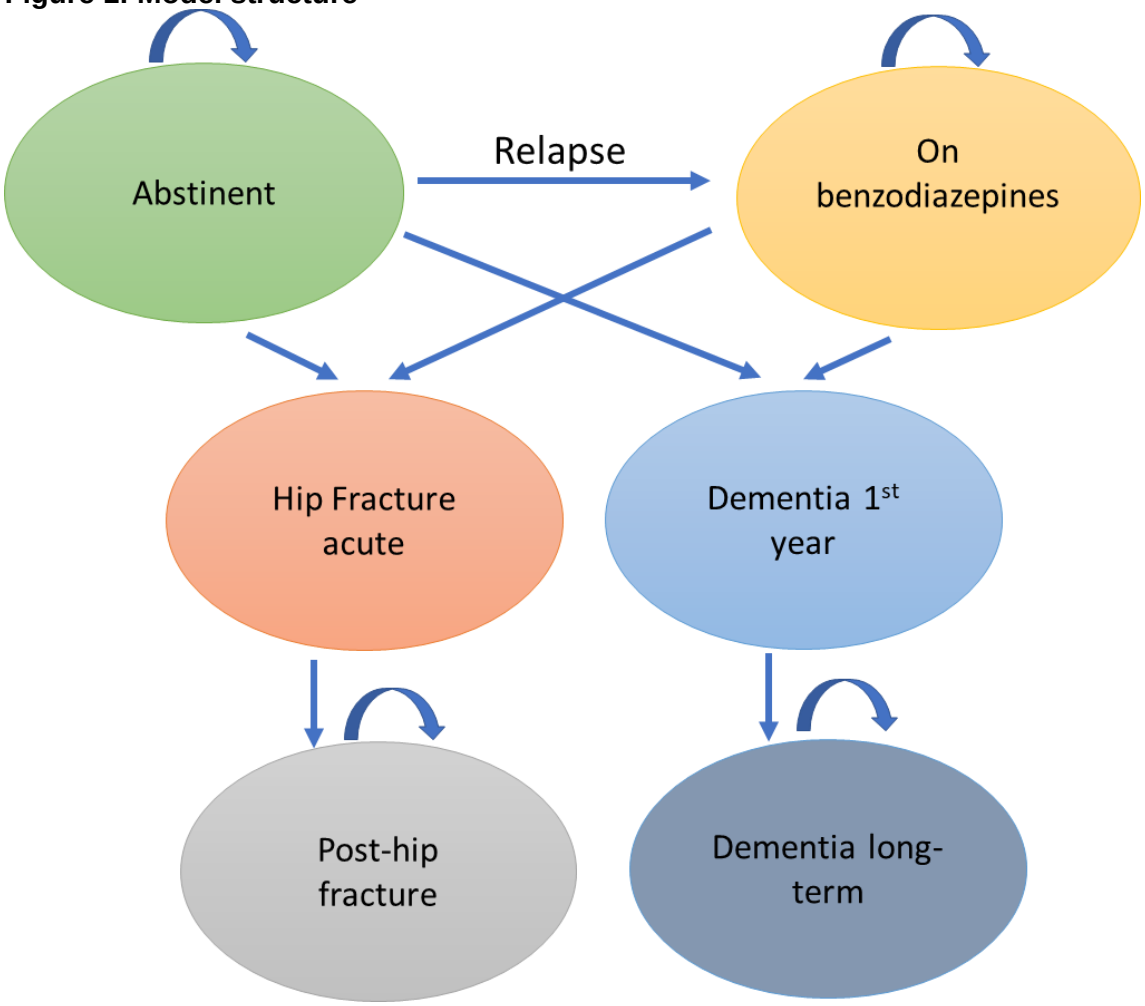
As mentioned before, the remaining adverse events were not modelled as explicit Markov states, but their incidence was used to calculate the number of events occurring during each cycle based on the population at risk. This is because they cause costs and utility losses only in the short-term, which can be captured without the need of using a different Markov state. Fatal events, such as suicides and fatal road traffic accidents, were incorporating into the transition to dead from all states by adding their excess mortality to the general population mortality (see sections 2.3.5.4). The increased risk of suicide caused by benzodiazepines was incorporated only in the sensitivity analysis (see section 2.5.5).

Dead is an absorbing state.

Summary of key assumptions:

- Relapses can occur only during the first two cycles of the model and once a person is abstinent at 2 years after the intervention, they will remain as such for the rest of their life
- People experiencing hip fracture or dementia immediately withdraw from benzodiazepines (if they were taking it) and cannot relapse. This was considered plausible by the committee as prescribing benzodiazepines or Z-drugs to people with poor mobility or dementia is commonly avoided, though there are cases where this still occurs (see section 4.2).
- Hip fracture can only be caused by a fall (this is to avoid double-counting error).
- Death following a fall injury can occur only as a complication of a hip fracture.
- People can experience either a hip fracture or dementia. There is no overlapping state.

Figure 1: Model structure



Note: people in each health state also have a state- and age-specific probability of transitioning to the dead health state.

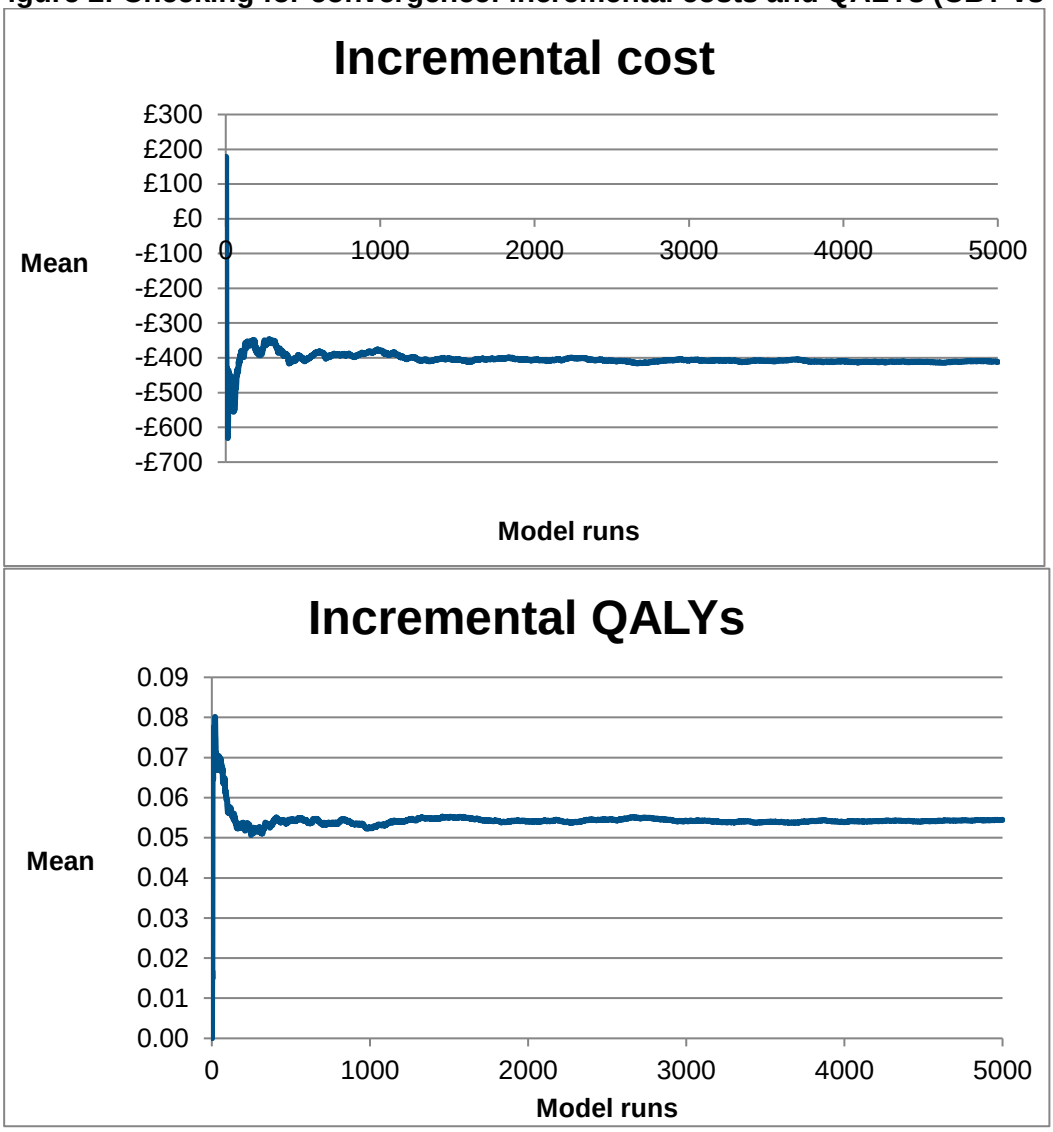
2.2.2 Uncertainty

The model was built probabilistically to take account of the uncertainty around input parameter point estimates. A probability distribution was defined for each model input

parameter. When the model was run, a value for each input was randomly selected simultaneously from its respective probability distribution; mean costs and mean QALYs were calculated using these values. The model was run repeatedly 10,000 times and results were summarised.

When running the probabilistic analysis, multiple runs are required to take into account random variation in sampling. To ensure the number of model runs was sufficient in the probabilistic analysis we checked for convergence in the incremental costs, QALYs for CBT vs TOA. This was done by plotting the number of runs against the mean outcome at that point (see example in Figure 2) for the base-case analysis. Convergence was assessed visually, and all had stabilised before 5000 runs.

Figure 2: Checking for convergence: incremental costs and QALYs (CBT vs TOA)



The way in which distributions are defined reflects the nature of the data, so for example event probabilities were given a beta distribution, which is bounded by 0 and 1, reflecting that the probability of an event occurring cannot be less than 0 or greater than 1. All of the variables that were probabilistic in the model and their distributional parameters are detailed in Table 1 and in the relevant input summary tables in section 2.3.1. Probability distributions in the analysis were parameterised using error estimates from data sources.

Table 1: Description of the type and properties of distributions used in the probabilistic sensitivity analysis

Parameter	Type of distribution	Properties of distribution
Hip fracture probability Baseline cessation probability	Beta	Bounded between 0 and 1. As the sample size and the number of events were specified alpha and beta values were calculated as follows: • Alpha = (number of patients hospitalised) • Beta = (number of patients) − (number of patients hospitalised)
Odds ratios Hazard ratios Risk ratios Incidence rates Incidence rate ratios	Lognormal	The natural log of the mean and standard error was calculated as follows: • Mean = ln(mean cost) − $SE^2/2$ • SE = [ln(upper 95% CI) − ln(lower 95% CI)]/(1.96×2) $\sqrt{\ln \frac{SE^2 + mean^2}{mean^2}}$ This formula includes a correction to ensure the mean generated in the probabilistic analysis will be the same as the reported mean.[6]
Utilities	Beta	Bounded between 0 and 1. Derived from mean and its standard error, using the method of moments. Alpha and Beta values were calculated as follows: Alpha = $mean^2 \times [(1-mean)/SE^2] - mean$ Beta = alpha×[(1−mean)/mean]
Utility decrements / Costs	Gamma	Bounded at 0, positively skewed. Derived from mean and its standard error. Alpha and beta values were calculated as follows: • Alpha = $(mean/SE)^2$ • Beta = SE^2/Mean

Abbreviations: 95% CI = 95% confidence interval; SE = standard error; SMR = standardised mortality ratio.

The following variables were left deterministic (that is, they were not varied in the probabilistic analysis):

- the cost-effectiveness threshold
- Health state and events costs (based on analyses that use unit costs from UK national sources)
- Drug costs (based on drug tariff which is known)
- Mortality probabilities for general population (based on UK national data)
- Utility score in the general population (based on the algorithm from Ara 2010[2])
- Prevalence data (based on population or large cohort studies)
- Initial cohort settings (based on a meta-analysis of the clinical trials included)

In addition, various deterministic sensitivity analyses were undertaken to test the robustness of model assumptions. In these, one or more inputs were changed, and the analysis was rerun to evaluate the impact on results and whether conclusions on which intervention should be recommended would change. Details of the sensitivity analyses undertaken can be found in methods section 2.5 Sensitivity analyses.

2.3 Model inputs

2.3.1 Summary table of model inputs

Model inputs were based on clinical evidence identified in the systematic review undertaken for the guideline, supplemented by additional data sources as required. Model inputs were validated with clinical members of the guideline committee. A summary of the model inputs used in the base-case (primary) analysis is provided in the table below. More details about sources, calculations and rationale for selection can be found in the sections following this summary table.

Table 2: Overview of parameters and parameter distributions used in the model

Input	Data	Source	Probability distribution
Comparators	• Tapering off plus cognitive behavioural therapy (CBT + TO) • Tapering off alone (TOA) • Usual care (UC)		n/a
Population	People continuously taking benzodiazepines		n/a
Perspective	UK NHS & PSS	NICE reference case[38]	n/a
Time horizon	Lifetime		n/a
Discount rate	Costs: 3.5% Outcomes: 3.5%	NICE reference case[38]	n/a
Cohort settings			
Cohort size	1,000		n/a
Cohort start age	64 years	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]	n/a
Percentage of females entering the model	62%	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]	n/a
Incidence of adverse events in the general population			
Falls annual incidence rate (population average)	Women = 0.505 Men = 0.392	HSE 2005[24]	Lognormal Women: LnR = -0.94 SE = 0.03 Men: LnR = -0.68 SE = 0.02
Falls - incidence rate ratio by age, women	65-69 = 1.000 70-74 = 1.036 75-79 = 1.096 80-84 = 1.518 85+ = 1.855	HSE 2005[24]	Lognormal 70-74: LnIRR = 0.04 SE = 0.09 75-79: LnIRR = 0.07 SE = 0.09 80-84: LnIRR = 0.42 SE = 0.06 85+: LnIRR = 0.62 SE = 0.05
Falls – incidence rate ratio by age, men	65-69 = 1.000 70-74 = 1.271 75-79 = 1.136	HSE 2005[24]	Lognormal 70-74: LnIRR = 0.24 SE = 0.08 75-79: LnIRR = 0.13 SE = 0.10

Input	Data	Source	Probability distribution
	80-84 = 1.644 85+ = 2.814		80-84: LnIRR = 0.50 SE = 0.07 85+: LnIRR = 1.03 SE = 0.04
Hip fracture annual probability, women	51-60 = 0.0002 61-70 = 0.0008 71-80 = 0.0037 81-90 = 0.0124 91+ = 0.0235	Hopkins 2012[25]	Beta 51-60: Alpha = 556 Beta = 2,302,261 61-70: Alpha = 1,281 Beta = 1,524,775 71-80: Alpha = 3,829 Beta = 1,025,096 81-90: Alpha = 7,424 Beta = 589,482 91+: Alpha = 2,652 Beta = 110,211
Hip fracture annual probability, men	51-60 = 0.0002 61-70 = 0.0006 71-80 = 0.0020 81-90 = 0.0063 91+ = 0.0160	Hopkins 2012[25]	Beta 51-60: Alpha = 521 Beta = 2,253,813 61-70: Alpha = 835 Beta = 1,444,646 71-80: Alpha = 3,829 Beta = 1,025,096 81-90: Alpha = 7,424 Beta = 589,482 91+: Alpha = 2,652 Beta = 110,211
Dementia annual incidence rate, women	65-69 = 0.0046 70-74 = 0.0064 75-79 = 0.0161 80-84 = 0.0396 85+ = 0.0553	CFAS II[30]	Lognormal 65-69: LnIRR = -5.38 SE = 0.38 70-74: LnIRR = -5.05 SE = 0.34 75-79: LnIRR = -4.13 SE = 0.24 80-84: LnIRR = -3.23 SE = 0.16 85+: LnIRR = -2.89 SE = 0.18
Dementia annual incidence rate, men	65-69 = 0.0050 70-74 = 0.0087 75-79 = 0.0167 80-84 = 0.0248 85+ = 0.0380	CFAS II[30]	Lognormal 65-69: LnIRR = -5.30 SE = 0.36 70-74: LnIRR = -4.74 SE = 0.28 75-79: LnIRR = -4.09 SE = 0.24 80-84: LnIRR = -3.70 SE = 0.24 85+: LnIRR = -3.27 SE = 0.27
RTA rate per person, women	Age-specific	Department of Transport[18]	n/a
RTA rate per person, men	Age-specific	Department of Transport[18]	n/a
Proportion of deaths caused by suicide, women	30-34 = 9.09% 35-39 = 5.8% 40-44 = 3.79% 45-49 = 2.86% 50-54 = 1.94% 55-59 = 1.02% 60-64 = 0.61% 65-69 = 0.39% 70-74 = 0.21% 75-79 = 0.14% 80-84 = 0.06%	Death Registrations Summary Statistics, England and Wales, 2017[45]	n/a

Input	Data	Source	Probability distribution
	85-89 = 0.03% 90-94 = 0.02% 95> = 0.01%		
Proportion of deaths caused by suicide, men	30-34 = 17.28% 35-39 = 12.21% 40-44 = 9.86% 45-49 = 7.53% 50-54 = 4.7% 55-59 = 2.5% 60-64 = 1.31% 65-69 = 0.68% 70-74 = 0.39% 75-79 = 0.25% 80-84 = 0.11% 85-89 = 0.12% 90-94 = 0.07% 95> = 0.01%	Death Registrations Summary Statistics, England and Wales, 2017[45]	n/a
Prevalence of health states in the general UK population			
Hip fracture prevalence	30-65 = 0 65> = 0.0088	Age UK 2019[1]	n/a
Dementia prevalence	30-60 = 0 60-64 = 0.9% 65-69 = 1.7% 70-74 = 3% 75-79 = 6% 80-84 = 11.1% 85-89 = 18.3% 90-94 = 29.9% 95> = 41.1%	Alzheimer's Society 2014[49]	n/a
Dementia severity	Mild 65-74 = 17.92% 75-84 = 18.97% 85>= 14.7% Moderate 65-74 = 40.83% 75-84 = 40.18% 85> = 31.53% Severe 65-74 = 41.25% 75-84 = 40.84% 85> = 53.78%	Wittenberg 2019[54]	n/a
BZD relative effects on adverse events			
BZD on falls incidence rate ratio	1.32	Richardson 2015[52]	Lognormal LnIR = 0.278 SE = 0.111
BZD on hip fracture risk ratio	1.52	Donnelly 2017[20]	Lognormal LnRR = 0.419 SE = 0.083
BZD on dementia OR	1.38	Lucchetta 2018[29]	Lognormal LnOR = 0.322 SE = 0.128
BZD on RTA OR	All ages = 1.6	Barbone 1998[5]	Lognormal

Input	Data	Source	Probability distribution
	<30 = 2.66 30-44 = 2.18 45-64 = 1.48 >65 = 1		All ages: LnOR = 0.47 SE = 0.14 <30: LnOR = 0.98 SE = 0.35 30-44: LnOR = 0.78 SE = 0.26 45-64: LnOR = 0.39 SE = 0.22
BZR on suicide OR	1.89	Cato 2019[9]	Lognormal LnOR = 0.637 SE = 0.243
Cessation, reduction and relapse			
Cessation probability with TOA	Cycle 0 = 0.526 Cycle 1 = 0.381	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]	Beta Cycle 0: Alpha = 60 Beta = 54 Cycle 1: Alpha = 45 Beta = 73
Dose reduction probability with TOA (only in SA)	Cycle 0 = 0.5 Cycle 1 = 0.364	Baillargeon 2003[4]	Beta Cycle 0: Alpha = 9 Beta = 9 Cycle 1: Alpha = 8 Beta = 14
Dose reduction probability with CBT (only in SA)	Cycle 0 = 0.875 Cycle 1 = 0.3	Baillargeon 2003[4]	Beta Cycle 0: Alpha = 7 Beta = 1 Cycle 1: Alpha = 3 Beta = 7
Freedom from relapses	TOA 1 year = 0.85 2 years = 0.69 CBT + TO 1 year = 0.71 2 years = 0.67	Morin 2005[34]	Beta TOA 1 year: Alpha = 11 Beta = 2 2 years: Alpha = 9 Beta = 4 CBT + TO 1 year: Alpha = 15 Beta = 6 2 years: Alpha = 14 Beta = 7
Intervention effectiveness			
CBT+TO vs TOA cessation risk ratio	Cycle 0 = 1.46 Cycle 1 = 1.30	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]	Lognormal Cycle 0: LnRR = 0.38 SE = 0.26 Cycle 1: LnRR = 0.26 SE = 0.33
TOA vs UC cessation risk ratio	Cycle 0 = 4.91	Oude Voshaar 2006[47]	Lognormal LnRR = 1.43 SE = 0.43
Mortality			
General population mortality	Age and sex dependent	ONS Life Tables 2016-2018[46]	n/a
Hip fracture mortality hazard ratio	Women= 3.7 Men= 2.87	Haentjens 2010[23]	Lognormal Women: LnHR = 1.31 SE = 0.06 Men: LnOR = 1.05 SE = 0.07
Post-hip fracture mortality hazard ratio	Women= 1.86 Men= 1.9	Haentjens 2010[23]	Lognormal Women: LnHR = 0.62 SE = 0.08 Men: LnOR = 0.64 SE = 0.10
Dementia mortality risk ratio	1.71	Rao 2016[51]	Lognormal LnRR = 0.54 SE = 0.15
Health-related quality of life (utilities)			
General population utility score	Age and sex dependent	Ara & Brazier[2]	n/a
CBT+TO utility score (18 months)	0.82	Oude Voshaar 2006[48]	Beta Alpha = 140.47 Beta = 30.97
TOA utility score (18 months)	0.79	Oude Voshaar 2006[48]	Beta Alpha = 69.9 Beta = 18.06

Input	Data	Source	Probability distribution
UC utility score (18 months)	0.83	Oude Voshaar 2006[48]	Beta Alpha = 46.22 Beta = 9.65
Dementia utility score	Mild = 0.74 Moderate = 0.79 Severe = 0.83	Landeiro 2020[27]	Gamma Mild: Alpha = 85.84 Beta = 0.003 Mod: Alpha = 4.96 Beta = 0.03 Sev: Alpha = 4.51 Beta = 0.05
Hip fracture disutility	0.237	Griffin 2015[22]	Gamma Alpha = 10.2 Beta = 0.02
Post-hip fracture disutility	0.22	Griffin 2015[22]	Gamma Alpha = 8.8 Beta = 0.02
Emergency admission disutility	0.014	Church 2012[10]	Gamma Alpha = 8 Beta = 0.002
Hospitalisation utility detriment	0.144	Church 2012[10]	Gamma Alpha = 0.5 Beta = 0.31
Admission to residence care utility detriment	0.06	Church 2012[10]	Gamma Alpha = 0.1 Beta = 1.07
RTA utility detriment	0.21	Church 2012[10]	Gamma Alpha = 25 Beta = 0.008
Costs			
CBT+TO cost	£438	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]	n/a
TOA cost	£281	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]	n/a
BZD pharmaceutical cost (annual)	£78	BNF[8] Prescription Cost Analysis[42]	n/a
RTA cost	Slight = £1,095 Serious = £15,719 Fatal = £6,166	Department of Transport[19]	n/a
Cost of a fall (excluding hip fracture)	All fall = £1,199 Serious fall = £5,995	NHS Reference Cost 2018-2019[43] PSSRU 2020[12] Craig 2013[11]	n/a
Hip fracture cost 1st year	£14,971	Leal 2016[28]	Gamma Alpha = 3,117.8 Beta = 4.802
Post-hip fracture cost (annual year 2+)	£2,260	Leal 2016[28]	Gamma Alpha = 571,2 Beta = 3.96
Cost of dementia 1st year	Mild = £3,478 Moderate = £4,457 Severe = £7,247	Wittenmberg 2019[54]	n/a

Input	Data	Source	Probability distribution
Cost of dementia long-term (annual year 2+))	Mild = £4,053 Moderate = £6,026 Severe = £7,844	Wittenmberg 2019[54]	n/a

Abbreviations: CBT+TO = Tapering off plus cognitive behavioural therapy; TOA = Tapering off alone; UC = Usual Care; BZD = Benzodiazepines; RTA = Road Traffic Accident; OR = Odds ratio.

2.3.2 Initial cohort settings

In the base case scenario, the initial characteristics of people entering the model (age and gender) were estimated using descriptive statistics data from the trials included in the clinical review. A weighted average of age and gender proportion was calculated using the sample size of the trials as weights and the results were used in the base case scenario (see Table 3). A sensitivity analysis on two different age groups (50 and 70 years old) was conducted to see the impact of the interventions on younger and older cohorts.

Table 3: Initial cohort settings in the base case scenario

Cohort setting	Value	Source
Mean age	64	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]
Proportion of female	62%	Baillargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]

2.3.3 Incidence of adverse events

Key adverse events related to benzodiazepine use were incorporated into the model. It was assumed that people who had withdrawn from benzodiazepines would experience these events at the same rate as the general population. The incidence of adverse events in the general population was identified through a non-systematic literature review of the available evidence, prioritising UK studies with large sample sizes. The studies ultimately selected to inform baseline incidence are presented in table 3 together with their characteristics and source.

Table 4: Sources for baseline incidence of adverse events

Event	Source	Country	Sample size
Falls	Health Survey of England 2005[24]	UK	2,673
Hip fractures	Hopkins 2012[25]	Canada	21,687
Dementia	CFAS II[30]	UK	5,288
Road traffic accidents	Department of transport[18]	UK	61,549
Suicides	ONS, Death registrations summary tables[44]	UK	3,930

Benzodiazepine relative effects were applied to the baseline incidence to obtain the incidence in a benzodiazepine user population (see section 2.3.4).

2.3.3.1 Falls

Falls incidence in the general population was taken from the Health Survey of England (HSE) of 2005 previously used and analysed for a health economics analysis[40] conducted alongside the NICE guideline CG161 on falls in older people[39] as no more recent data was found. Results from the study on a subpopulation of older people (>65) are reported elsewhere.[24] Overall, the HSE gives an annual rate of 0.46 fall per person with a higher rate in women than men in people older than 65 as shown in table 5. This rate is higher than the fall prevalence (26%) as some people fall multiple times during a year. Overall, this rate is similar to the rates estimated by other published studies which estimated a fall rate ranging between 0.41 to 0.54,[50] but it is more recent and therefore it was chosen for this economic analysis.

Table 5: Fall incidence parameters [a]

Parameter	Women	Men	Source
Fall rate per person per year (>65)	0.505 (0.484 to 0.525)	0.392 (0.372 to 0.412)	HSE 2005[24]
Incidence rate ratio by age group			
65-69	1	1	HSE 2005[24]
70-74	1.036 (0.863 to 1.21)	1.271 (1.065 to 1.477)	
75-79	1.069 (0.891 to 1.247)	1.136 (0.909 to 1.363)	
80-84	1.518 (1.345 to 1.69(	1.644 (1.408 to 1.88)	
85+	1.855 (1.677 to 2.033)	2.814 (2.585 to 3.043)	
Proportion of people by age			
65-69	27%	32%	HSE 2005[24]
70-74	24%	27%	
75-79	21%	20%	
80-84	17%	13%	
85+	12%	8%	

(a) CIs have been calculated using reported SEs

Data on falls for people younger than 65 was not available. As the model includes people aged 50 and over, fall incidence rates for people aged between 50 to 65 had to be extrapolated. For people aged 61 to 65 we assumed the same fall rate of the age group 65-69. For people aged 50 to 60, the same assumption could not be used as the fall rate is expected to be considerably lower in this population. As hip fractures are often caused by falls, we expected hip fractures and falls to follow similar trends by age. Data on hip fractures for people aged 51 to 60 were available (see section 2.3.3.2) and were used to extrapolate the fall rate in this group. For this purpose, a multiplier factor was calculated by dividing the hip fracture probability in people aged 51-60 by the probability in the 61-70 age group. This factor was then applied to the incidence rate of falls of people aged 65-70 to calculate the incidence rate of falling in the age group 51-60.

2.3.3.2 Hip fracture

The annual probability of experiencing a hip fracture was taken from a Canadian study including 21,687 fractures which occurred in people aged 50 or older. The study, albeit not British, was chosen for this economic analysis as the Canadian setting was considered relatively similar to the UK and no UK data was available. Probabilities by age and gender are presented in Table 6.

Table 6: Hip fracture and annual probability

Age group	Women	Men	Source
51-60	0.0002	0.0002	Hopkins 2012[25]
61-70	0.0008	0.0006	
71-80	0.0037	0.0020	
81-90	0.0124	0.0063	
91+	0.0235	0.0160	

Hip fracture probability and fall incidence rates were used to calculate the number of hip fractures and falls occurring at each cycle. The model takes into consideration the proportion of men and women alive at the beginning of each cycle to calculate a weighted average incidence rate which is applied to the population at risk. As the events can occur an any point during the cycles and not necessarily at the end or the beginning, a half-cycle correction was built into the analysis. One of the key assumptions of the model is that a hip fracture can only be the consequence of a fall. Hence, at each cycle, the number of hip fractures predicted is subtracted from the number of falls to avoid double counting.

2.3.3.3 Dementia

The incidence of new cases of dementia was taken from the MRC Cognitive Function and Ageing Study (CFAS II),[30] a recent British study with 5,288 participants interviewed over 3 years. The authors estimated incidence rates of dementia for six age groups split by gender (see Table 7).

Table 7: Dementia incidence rate per person per year

Age group	Women	Men	Source
65-69	0.005 (0.002 to 0.01)	0.005 (0.002 to 0.1)	CFAS II[30]
70-74	0.006 (0.003 to 0.012)	0.009 (0.005 to 0.015)	
75-79	0.016 (0.01 to 0.026)	0.017 (0.01 to 0.026)	
80-84	0.04 (0.03 to 0.05)	0.025 (0.015 to 0.04)	
85+	0.055 (0.039 to 0.078)	0.0380 (0.022 to 0.064)	

The incidence rates shown in Table 6 were transformed into probabilities using standard formulas and used to determine the age- and gender-specific probability that a person will develop dementia every year in the model. As illustrated in section 2.2.1, people who develop dementia move to a tunnel state and then to a long-term state where they are assumed to remain for the rest of their life.

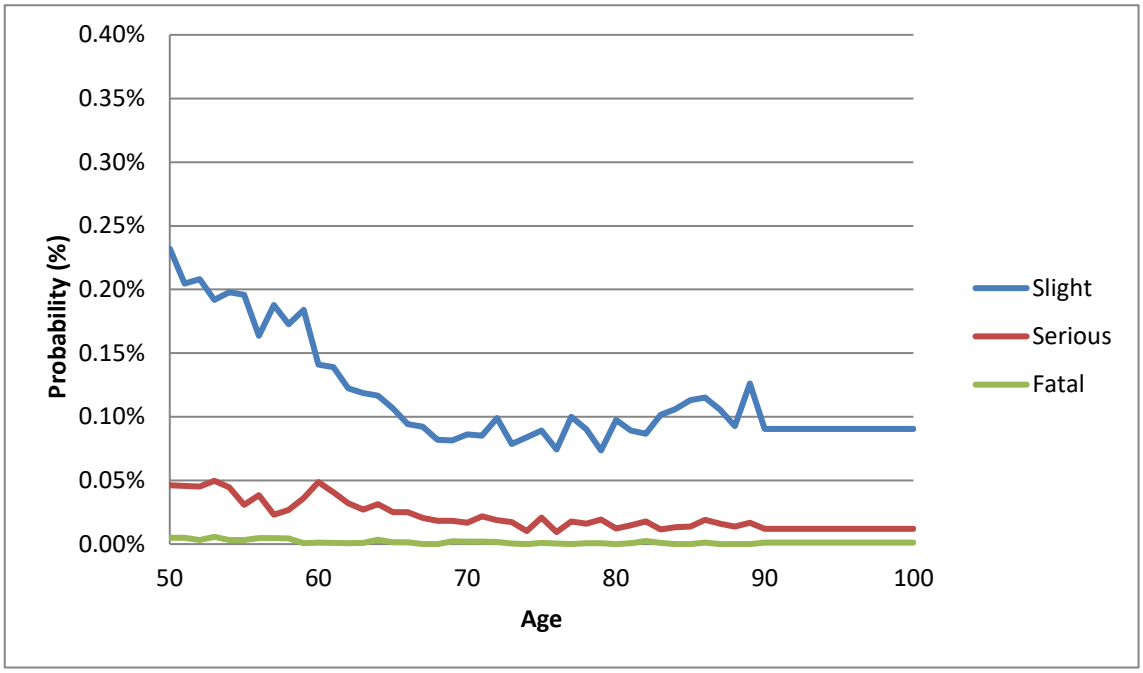
Dementia incidence rates of people younger than 65 could not be found. Prevalence data from the UK Alzheimer's society[49] shows that dementia is rather uncommon before age 65 and extremely rare before age 50 and, consequently, the model assumes that dementia can occur only from age 65 onward, with no early onset of dementia included. Early-onset of dementia (defined as age less than 65 years), although very rare, may be affected by benzodiazepine usage, and thus the model may underestimate some of the benefits of the intervention in terms of a reduction of cases of early-onset of dementia.

2.3.3.4 Road traffic accidents

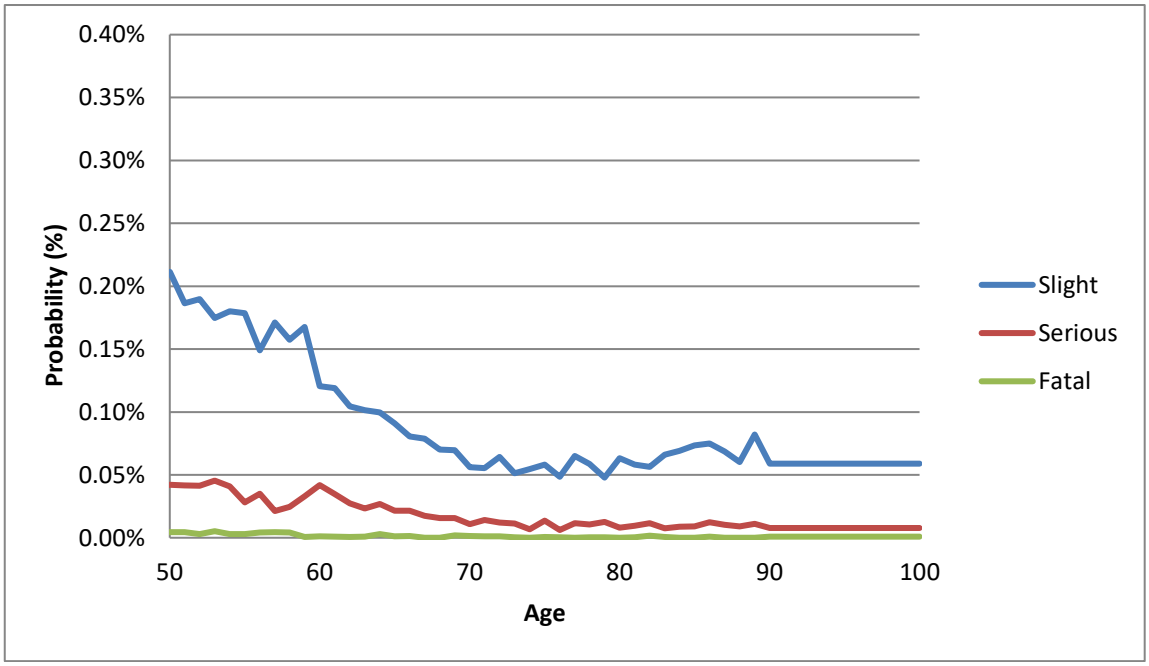
Department of Transport data were used to estimate the number of drivers in England split by age and gender. This was achieved by multiplying the age- and sex- specific probability of holding a licence by the corresponding population in England reported by the Office for National Statistics in 2019.

The number of road traffic casualties in England in 2019 was disaggregated according to age, sex and severity of the casualty (slight, serious or fatal). The total number of casualties was divided by the number of drivers in each age range and sex category to calculate the probability of slight, serious, and fatal RTAs for male and female drivers over a lifetime. This was then converted to a probability of slight, serious and fatal RTA in the overall population by multiplying by the proportion of the population that are drivers (Figure 3 and Figure 4).

Figure 3: Probability of a road traffic accident in males over 50 years in England, by age



`Figure 4: Probability of a road traffic accident in females over 50 years in England, by age



The gender-adjusted probability of having an accident was used to determine the number of accidents occurring at each cycle and to calculate the costs and loss of QALYs associated

with them. Fatal accidents always result in death. However, the life tables used to estimate general population mortality already include deaths due to traffic accidents, so adding fatal accident deaths at each cycle would have led to double counting and to an overall over-estimation of total deaths. Instead, the “excess mortality” or the additional mortality due to fatal road traffic accidents caused by benzodiazepine use was added at each cycle as this was not captured by the life tables (see chapter 2.3.5.4 for a further explanation).

2.3.3.5 Suicide

The incidence rates of suicide were calculated using the death summary tables from 2017,[44] reporting all death causes for people aged 0 to 100. Items X60-X84 are classified as intentional self-harm resulting in death and were considered as suicide in this analysis. The number of suicides in each age category was divided by the overall number of deaths to calculate the proportion of deaths caused by suicides. This number was then combined with the lifetables 2016-2018 to obtain the probability of dying due to suicide. At each cycle, the model used the probability of dying due to suicide to calculate the number of suicides occurring over the lifetime of the cohort.

2.3.3.6 Incidence of events in benzodiazepine users

Incidence rates in benzodiazepine users are higher than in the general population as several studies have shown a statistically significant correlation between benzodiazepine use and the events included in this model. To calculate the incidence in this population, several studies were reviewed, prioritising large meta-analyses with big sample size (see table 8).

Table 8: Relative effects of benzodiazepines use on adverse events

Event	Parameter	Value	Source
Falls	Incidence rate ratio	1.32 (1.05 to 1.65)	Richardson 2015[52]
Hip fracture	Risk ratio	1.52 (1.37 to 1.9)	Donnelly 2017[20]
Dementia	Odds ratio	1.38 (1.07 to 1.77)	Lucchetta 2018[29]
Road traffic accidents	Odds ratio	1.6 (1.2 to 2.1) (split by age in the sensitivity analysis)	Barbone 1998[5]
Suicide	Odds ratio	1.89 (only in the sensitivity analysis)	Cato 2019[9]

Falls were included in the model as incidence rates, so risk ratios or odds ratios could not be applied to calculate falls occurring in people taking benzodiazepines as these can be applied to probabilities only. Likewise, hazard ratios can be applied only to hazard rates but not to incidence rates. Therefore, a study based on The Irish Longitudinal study[52] on ageing enrolling 6,666 adults was used as it reported incidence rate ratio (IRR) instead of other relative effect ratios. The IRR is the ratio between the number of falls occurring in the benzodiazepine cohort and the number of falls occurring in a cohort not using benzodiazepines. This ratio was adjusted for baseline covariates: age, sex, living alone, education, employment status, income, smoking status, the time between interviews, co-morbidities, incontinence, pain, sleep problems, depressive symptoms, cognition, self-rated vision, self-rated hearing, disability, public health-care coverage, history of falls, fracture, fainting and hospitalisation. Hence, the adjusted IRR should capture the causal and real effect of benzodiazepine on falls rate, and was included in the model to calculate the number of falls occurring in the benzodiazepine cohort.

Hip fracture rates in people assuming benzodiazepines were estimated using a recent meta-analysis on 18 studies[20], none of them being a randomized controlled trial. Similar to falls,

the analysis found benzodiazepine use to increase the risk of incurring a hip fracture by around 50% with a risk ratio of 1.52.

Cases of dementia among benzodiazepine users were estimated using a meta-analysis from 2018[29] on the association between the development of dementia and the use of benzodiazepines. A total of 980,860 adults were included and the results of the meta-analysis suggest that benzodiazepine use is a risk factor for developing dementia with an odds ratio of 1.38.

A study on Lancet from 1998[5] on 19,386 British drivers found benzodiazepines to increase the risk of a road traffic accident, with the effect decreasing with driver's age. The overall hazard ratio of 1.6 was used in the base case analysis to estimate the number of road traffic accidents occurring in the benzodiazepine user population whereas, in the sensitivity analysis, the odds ratio was split by age group.

The association between suicide and benzodiazepine use is less certain. Although some studies suggested that benzodiazepines may lead to an increased risk of suicidal behavioural,[41] the causal mechanism is uncertain as part, if not all, of the causal effect of benzodiazepines, may be explained by reverse causality and selection bias. In other words, people taking benzodiazepines may be at a higher risk of suicide because of their underlying conditions and not because of the drugs they are taking. A propensity-matched score study[9] found benzodiazepine use to increase the risk of suicide with an odds ratio of 1.89. However, the study design used is insufficient to rule out reverse causality and selection bias. Therefore, the committee agreed to use the relative effect from this study only in the sensitivity analysis and assume no increased risk of suicide with benzodiazepines in the base case.

2.3.4 Intervention effectiveness

Data on the effectiveness of the three strategies considered were taken from the clinical review meta-analysis and used to determine the proportion of patients under benzodiazepine in each intervention.

2.3.4.1 Cessation post-intervention

Cessation rates in the TOA arm were pooled together across all the trials to calculate the probability of post intervention cessation after tapering alone, which was used as the baseline probability. It should be noted that the first follow-up after the intervention was done at different points in time in each trial. In Morin, post intervention follow-up was done 10 weeks after the end of the treatment, in Bailargeon 8 or 12 weeks later and in Morin around 13 weeks later. Therefore, the meta-analysis incorporated data collected at different points in time, although not so distant from each other to cause significant issues. The probabilities of post-intervention cessation in the CBT and usual care strategies were then calculated by applying the corresponding risk-ratios from the meta-analysis undertaken as part of the clinical review for this guideline to the probability of cessation with TOA. Usual care was included as a strategy only in one trial,[47] so its risk ratio was estimated from this single trial only. See Table 9 for the data used.

Table 9: Post-intervention cessation data

Parameter	Value	Source
TOA cessation rate probability (PI)	0.526	Bailargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]
CBT vs TOA risk ratio (PI)	1.46 (0.87 to 2.43)	Bailargeon 2003[4] Morin 2004[33]

Parameter	Value	Source
		Oude Voshaar 2006[47]
TOA vs UC risk ratio (PI)	4.19 (1.82 to 9.65)	Oude Voshaar 2006[47]

Abbreviations: CBT = Cognitive behavioural therapy; TOA = Tapering off alone; UC = Usual care.

These probabilities were used to determine in which proportion people start in each state at cycle 0 of the Markov model.

2.3.4.2 Relapse

As described in section 2.2.1, relapse or transit from abstinence to the relapsed state (On benzodiazepines) is allowed for the first 2 cycles, as the committee's clinical experience and published literature showed that, beyond 2 years, relapsing becomes relatively rare.[14] In the sensitivity analysis, this assumption was tested by exploring a scenario where relapses occur for five years (see Section 2.5.2).

Two different approaches were used to model relapses in the CBT and TOA arms and will be presented in this section. Relapse after usual care was shown not to occur in the single trial including usual care as a strategy[47] so the model allows relapses only in the other two arms. In a sensitivity analysis, the relapse rate of TOA was applied to the usual care asm as well.

Relapse between cycle 0 and cycle 1 was obtained using data relative to cessation rates at 1 year after the intervention (see Table 10).

Table 10: Cessation rate at 1 year

Parameter	Value	Source
Usual care cessation rate (1 year)	0.147	Oude Voshaar 2006[47]
TOA cessation rate (1 year)	0.381	Bailargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]
CBT vs TOA risk ratio (1 year)	1.3 (0.67 to 2.47)	Bailargeon 2003[4] Morin 2004[33] Oude Voshaar 2006[47]

Comparing cessation at 1 year (Table 10) to post-intervention cessation (Table 9), it can be seen that cessation after TOA and CBT is lower 1 year after the intervention, suggesting that some relapses occur between year 0 and year 1. The probability of relapse between year 0 and 1 in the CBT and TOA groups were therefore calculated using the following equation:

$$P_{Relapse} = \frac{P^0_{Cessation} - P^1_{Cessation}}{P^0_{Cessation}}$$

Where $P_{Relapse}$ is the probability of relapse between cycle 0 and 1, $P^0_{Cessation}$ is the probability of post-intervention cessation (cycle 0) and $P^1_{Cessation}$ is the probability of cessation at year 1. This equation was used to calculate the two relapse probabilities in the CBT and TOA arms (see Table 11).

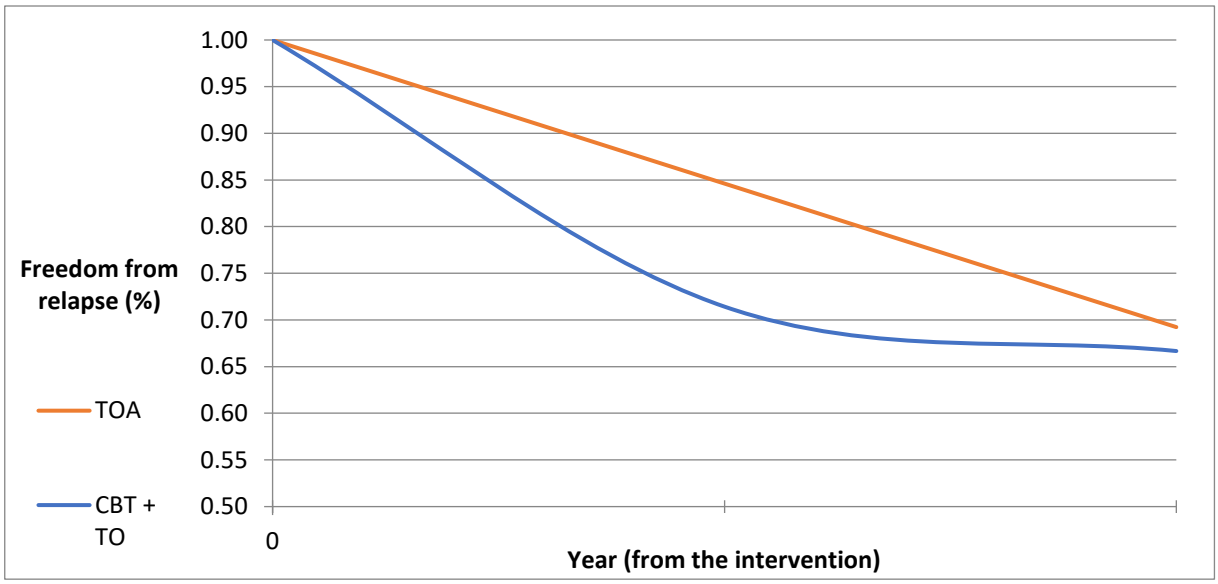
Table 11: Relapse probability between year 0 and 1

Parameter	Value
TOA relapse probability	0.27
CBT relapse probability	0.33

Although the committee thought that relapses should be considered at least until the second year, most of the trials in the clinical review do not have any follow-up data after the first year.

The only exception is the follow-up study of Morin 2005[34], who conducted a survival analysis on relapses occurring up to 2 years after the intervention on one of the RCTs included in the review.[33] The results of the survival analysis are illustrated in the Kaplan-Meier curves in **Figure 5.**

Figure 5: Relapse survival analysis from Morin 2005[34]



Source: Morin 2005[34]

The survival analysis showed that the combined strategy (CBT+TO) and TOA, although following different patterns during the first year, seem to converge at a similar relapse rate at 2 years. Data points from the Kaplan-Meier curves of CBT+TO and TOA were extracted and used to calculate the probability of relapsing between years 1 and 2 (Table 12).

Table 12: Relapse probability between year 1 and 2

Parameter	Value
TOA relapse probability	0.18
CBT relapse probability	0.07

This approach ensures that the relapse rate decreases between the first and the second year, which is generally expected, as people who remained abstinent for one year are less likely to relapse in the second year. On the other hand, the low sample size of the survival analysis (13 for TOA and 21 for CBT+TO) may add uncertainty to the analysis and, as a result, another scenario was tested in the sensitivity analysis where the relapse rates at the first year (Table 11) were applied in the second cycle as well (see section 2.5).

2.3.5 Mortality

2.3.5.1 General population

Population mortality by gender was based on data from life tables for England 2018-2019.[46] Cycle and gender-specific mortality vectors were calculated taking into account age and differences in mortality between gender. Additional mortality caused by the health states and events (hip fracture, dementia, road traffic accidents and suicide) was calculated separately for females and males by using gender-specific mortality relative risks (see next sections). At each cycle, the model uses gender split, which can change over time as men die at a higher rate than women, to calculate the gender-weighted average probability of dying in each state, which is used to determine the overall mortality of the cohort in each cycle. As population

mortality is not available beyond 100 years, the model applied the mortality rate for age 100 to those who are 100 years or older.

2.3.5.2 Mortality with dementia and hip fracture

Mortality in the population with dementia and hip fracture was calculated using the relative effects shown in Table 13.

Table 13: Mortality relative effects

Event	Parameter	Value	Source
Dementia	Risk ratio`	1.71 (1.27 to 2.29)	Rao 2016[51]
Hip fracture	Hazard ratio	Males: 3.7 (3.31 to 4.14) Females: 2.87 (2.52 to 3.27)	Haentjens 2010[23]
Post hip fracture	Hazard ratio	Males: 1.90 (1.58 to 2.3) Females: 1.86 (1.6 to 2.16)	Haentjens 2010[23]

The dementia risk ratio was taken from a meta-analysis of hospital administrative database studies including 11 studies conducted within the last 15 years and consisting of outcomes from 1,044,131 dementia patients compared to 9,639,027 elderly patients without dementia. Most of the studies were conducted in the US and none was done in the UK. The analysis found patients with dementia to have a higher mortality than patients without dementia with a risk ratio of 1.71. This factor was applied to females and males alike as data were not stratified by gender.

A risk ratio or relative risk is defined as the ratio of the probability of an event occurring in the exposed group versus the probability of the event occurring in the non-exposed group. Therefore, this risk ratio could not be directly applied to the general population mortality (see 2.3.5.1) as the latter refers to the actual English population, a group consisting of both exposed and non-exposed individuals. Failing to recognize this would inevitably lead to an overestimation of the overall mortality by the model. Therefore, the mortality in the non-exposed group had to be calculated first using dementia prevalence data and the following equation which is an adaptation from equation 3.6 at page 49 of the Applied Methods of Cost-Effectiveness Analysis in Health Care handbook:[21]

$$P_{Non-exposed} = \frac{P_{general\ population}}{(1-x) + xRR}$$

where $P_{general\ population}$ is the probability of dying in the general population (calculated using the ONS life tables), x is the prevalence of the disease (dementia) and RR is the mortality risk ratio of dementia. Once the probability of dying in the non-exposed group was obtained, the probability of dying in people with dementia could be easily estimated by multiplying the risk ratio from Table 11 to $P_{Non-exposed}$. This was done for each year of age within each group to obtain an age-specific vector of transition probabilities.

The hazard ratios for hip fracture and post-hip fracture states were taken from a meta-analysis[23] looking at excess mortality of people with hip fracture compared to the general population up to 10 years after a fracture. The studies included in the meta-analysis are mostly conducted in Europe (14 out of 24) although only one was conducted in the UK and were published during a period ranging from 1979 to 2009. The hazard ratio in the first year was used for the hip fracture state, as this state represents the acute phase of the condition. For the post-hip fracture state, it was decided to use the hazard ratio for years 1-2 as this was found to be relatively similar to the hazard ratios reported for the following years. As this

analysis looked at mortality outcomes in people with hip fracture up to 10 years only, it is possible that beyond the last follow-up the hazard ratio decreases or approaches 1. However, the trend seems to suggest that the hazard ratio remains stably above one over time and, for this reason, we assumed that post-hip fracture state is a permanent state, where people remain until they die (see section 4.2 for further discussion).

The authors stated that mortality in the reference group "was based on 2004 U.S. life-tables published by the National Center for Health Statistics". Hence, differently from the dementia risk ratio seen before, this hazard ratio is the ratio between mortality in the exposed group and mortality in the general population (including exposed and non-exposed individuals). For this reason, the equation used for dementia was not necessary to determine the probability of dying in the non-exposed group, and mortality in people with hip fracture was simply calculated by multiplying the hazard ratio in table 13 by the mortality rate in the general population. This was done for each year of age to obtain an age-specific vector of rates for both hip fracture and post-hip fracture states. Rates were then converted into transition probabilities using standard formulas and were applied to the gender-specific mortality vectors described in section 2.3.5.1.

2.3.5.3 Mortality in relapsed and abstinent states

Once mortality in people with dementia and hip fracture was known, mortality in the two states with no health condition, namely relapsed and abstinent, had to be estimated as well. This differs from mortality in the general population as this latter group includes also people who have dementia or hip fracture whereas, in the relapsed and abstinent states, nobody has neither of these two conditions. Therefore, the probability of dying in these two states was calculated using the following equation:

$$P_{Stable} = \frac{P_{general\ population} - x_1\ P_1\ - x_2\ P_2}{(1 - x_1 - x_2)}$$

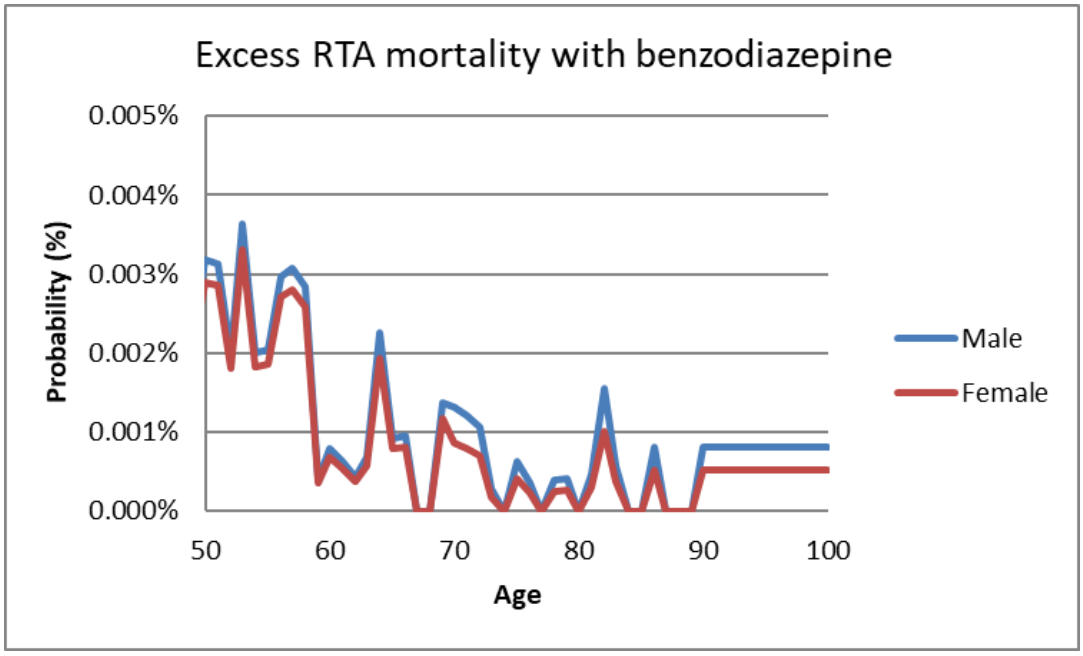
where x_1 and x_2 are, respectively, the prevalence of dementia and hip fractures in England and P_1 and P_2 are the probabilities of dying in people with dementia and hip fracture calculated through the approach described in 2.3.5.

Mortality in the stable state was calculated for each year of age to obtain an age-specific vector of transition probabilities which was applied to people in both abstinent and relapsed states although the excess mortality due to RTA was added to the mortality of those under benzodiazepines (see next section).

2.3.5.4 Road traffic accident - excess mortality with benzodiazepine use

As discussed in chapter 2.3.3.4, data from the Department of Transport were used to determine the probability of being victim of a road traffic accident. Fatal accidents always result in death. However, these deaths could not be added to the model as the life tables used to inform general population mortality already include deaths caused by road traffic accidents. The excess mortality, or the additional RTA mortality caused by benzodiazepines in those who are in the benzodiazepine state (captured by using the overall population odds ratio in the base case), is not captured by the life tables and so had to be added to the model. To estimate the excess mortality, the probability of dying because of a fatal RTA in the general population was subtracted from the probability of being the victim of a fatal RTA if taking benzodiazepines. This latter was calculated by applying the odds ratio of 1.6 shown in table 8 of section 2.3.3.6 to the probability of experiencing a fatal road traffic accident in the general population. The benzodiazepine excess mortality for both males and females are illustrated in **Figure 6** and represent the additional risk of dying of an RTA in people who are taking benzodiazepines.

Figure 6: Excess road traffic accident mortality with benzodiazepines



These probabilities were multiplied for the number of people in the relapsed state at each cycle of the model to determine the additional number of RTA deaths caused by benzodiazepines during the lifetime of the cohorts.

2.3.6 Utilities

Utility scores in the relapsed and abstinent states in the three arms were taken from one of the randomized trials included in the clinical review[48] (see section 2.3.6.1).

Utility in the other states (dementia and hip fracture) as well as utility decrements caused by falls and road traffic accidents were identified through a non-systematic literature review of available published evidence (see sections 2.3.6.2 and 2.3.6.3).

2.3.6.1 Utility with CBT+TO, TOA and UC

No studies reporting EQ-5D could be identified in the clinical review or through a quality-of-life systematic search. One of the trials[48] reported the SF-36 scores at baseline and 18 months after the interventions: CBT+TO, TOA and usual care (see appendix B). SF-36 data seem to suggest that there is a slight QoL advantage in people in usual care, followed by people in CBT+TO arm and lastly people in TOA arm. It was noted that this trial failed to find CBT+TO compared with TOA effective. Therefore, it may be possible that the trial is overestimating quality of life in the CBT group, as this, according to the meta-analysis, should have a larger number of people abstinent and, therefore, at risk of withdrawal symptoms.

The SF-36 scores were mapped into EQ-5D-3L (UK tariff) utility scores using the following mapping algorithm from Ara and Brazier (model 5):[2]

$$EQ-5D = -0.18105 + 0.00781PF + 0.00213SF + 0.00022RE + 0.00599MH + 0.00472BP + 0.00064GH - 0.00069Age - 0.00004PF * PF - 0.00001SF * SF - 0.00003MH * MH - 0.00001BP * BP$$

where PF is physical functioning, SF is social functioning, RE is role limitation (emotionally), MH is mental health, GH is general health and Age is the average age of the cohort which was 63 at baseline in the original trial.[48]

The resulting EQ-5D utility scores are presented in table 14. Utility at 18 months was corrected for baseline differences found before the intervention.

Table 14: EQ-5D utility scores in CBT+TO, TOA and UC arms

Intervention	Baseline (before the intervention)	18 months	18 months with baseline correction	Source
CBT+TO	0.78 (0.23)	0.82 (0.22)	0.82 (0.22)	Oude Voshaar 2006[48]
TOA	0.78 (0.23)	0.79 (0.33)	0.80 (0.33)	
UC	0.80 (0.24)	0.85 (0.26)	0.83 (0.26)	

As mentioned, even after correcting for baseline differences, people in the usual care arm showed the highest utility score followed by people in the CBT arm and finally by people in the tapering-off alone arm. The committee hypothesized that the higher utility score in patients in usual care may be due to the fact that these people are not experiencing long-term withdrawal symptoms as there was no attempt to discontinue the medication in this arm. In addition, people in the UC arm do not receive any pressure from the doctor to discontinue their medication, which may affect the psychological aspect of their quality of life. Conversely, people in TOA and CBT+TO arms followed a gradual discontinuation programme and are at risk of developing withdrawal symptoms which, in some cases, may last more than 12 months as confirmed by some members of the committee. The study reporting quality of life (Oude Voshaar 2006[48]) did not find any difference in effectiveness between CBT+TO and TOA, therefore the higher quality of life in the CBT+TO arm should only be caused by the long-term quality of life benefits of the CBT programme. The committee noted that Oude Voshaar was not consistent with the other trials which found CBT to be effective against TOA. Hence, it is possible that the utility score of people in the CBT arm used in this model does not reflect the real utility score of people undergoing this intervention, as the committee expect CBT to improve benzodiazepine cessation rate and, consequently, to increase the number of people at risk of experiencing withdrawal symptoms.

If the difference in utility between the two intervention arms and usual care is due to withdrawal symptoms, then this difference will likely converge to zero over time as the symptoms will diminish until they disappear. However, the committee could not find a consensus on when this was supposed to happen. Hence, we propose three scenarios.

In the first scenario, that was deemed more realistic and chosen for the base case analysis, utility in the three arms converge to the utility score in the usual care in the second cycle, 2 years after the intervention. In the second scenario, tested in the sensitivity analysis, convergence occurs only at the fifth cycle, 5 years after the intervention. Finally, in the last scenario, differences in utility score between the three arms are assumed to be permanent and to persist for the whole life of the cohort. This last scenario was considered the least realistic one, but nonetheless tested in the sensitivity analysis.

The utility scores from the trial were compared with the population utility norms reported in the trial itself to estimate a utility multiplier, representing the relative utility of people enrolled in the trial compared with the general population. This multiplier was then applied to the utility scores of the general UK population estimated through an algorithm by Ara and Brazier[3] at each year of age to calculate the age- and arm-specific utility score. This methodology ensured that utility decreases with ageing as expected in the real world.

As the trial collected the utility score in the whole arm, we were unable to assign a different utility score to people in the abstinent or on benzodiazepines state, as it would be preferred. However, the lower utility in the abstinent state due to the onset of withdrawal symptoms should still be indirectly captured with this approach as the arms with the higher numbers of abstinent people (TOA and CBT+TO) are also the arms with the lower utility score assigned.

2.3.6.2 Utility detriments after a fall

Falls cause a loss of utility lasting for a cycle only unless they cause a hip fracture (see section 2.3.6.3). Utility losses caused by a fall injury were identified from an existing and published model reporting an annual loss of utility caused by falls-related events and are shown in Table 15.

Table 15: Annual utility loss after a fall-related event

Event	Loss of EQ-5D[a]	Source
Emergency admission	-0.014 (-0.016 to -0.01)	Church 2012[10]
Hospitalisation	-0.144 (-0.255 to 0)	
Discharge to residential care	-0.060 (-0.338 to -0.03)	

The proportion of falls resulting in an emergency admission, hospitalisation and admission to residential care was calculated using data from NHS Scotland from a Scottish cost study[11] (see Table 16). This was also used to estimate the average cost of a fall in England (see section 2.3.7.3). Hospitalisation and emergency admission are not treated as independent events, meaning that if a person is admitted to the emergency room and is then hospitalised, he would incur only a loss of utility equal to 0.144. By contrast, being admitted to a residential care is treated as an independent event, causing an additional loss of utility, as done in the original model.[10]

Table 16: Falls-related events

Event	Proportion	Source
Serious fall	20% of total falls	Craig 2013[11]
A&E attendance	80% of serious falls	
Hospitalisation	35% of A&E attendance	
Discharge to residential care	36% of hospitalisations	

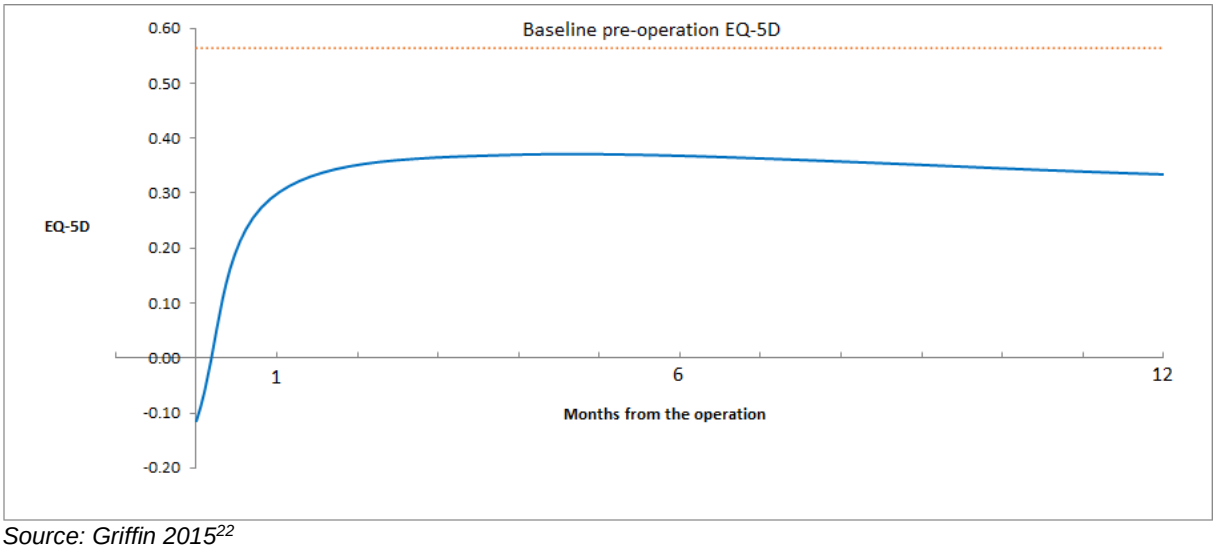
At each cycle, the model calculates the proportion of falls resulting in emergency admissions, hospitalisation and discharge to residential care and applies the utility decrements associated with each event. All the events are assumed to be transitory, hence the losses of utility occur for one cycle only. In the probabilistic analysis, these losses were allowed to vary independently.

2.3.6.3 Utility detriments after a hip fracture state

People in the hip fracture and post-hip fracture states are supposed to have a lower quality of life than the general population as a consequence of their reduced mobility and pain.

Griffin and colleagues using a prospective cohort of 741 patients treated at a single major trauma centre in the United Kingdom[22] found that people who experienced a hip fracture, have a reduction in EQ-5D (UK tariff) at 12 months of 0.22 (see **Figure 7**

Figure 7: Changes in EQ-5D following a hip fracture



Source: Griffin 2015[22]

As the graph shows, people start to improve immediately after the intervention but did not recover to baseline level at one year. Furthermore, the trend of the curve suggests that their quality of life may remain impaired in the long-term as people do not seem to improve much after four months.

To calculate the QALY loss during the first year after a hip fracture, the area above the curve EQ-5D in **Figure 7** was estimated using the trapezoidal rule and assuming linear change between data points. This gives a QALY loss of 0.237 that was assigned to people in the hip fracture state.

For the utility score in the post-hip fracture, it was assumed that the reduction in EQ-5D at 12 months of 0.22 was permanent. This assumption is not backed by data as the last follow-up of the study is at 12 days, though the dotted line in **Figure 7** appears to become flatter when reaching the last data point, although still slightly increasing. It is possible therefore that the model overestimates utility loss caused by a hip fracture, and this was mentioned as a limitation (see section 4.2).

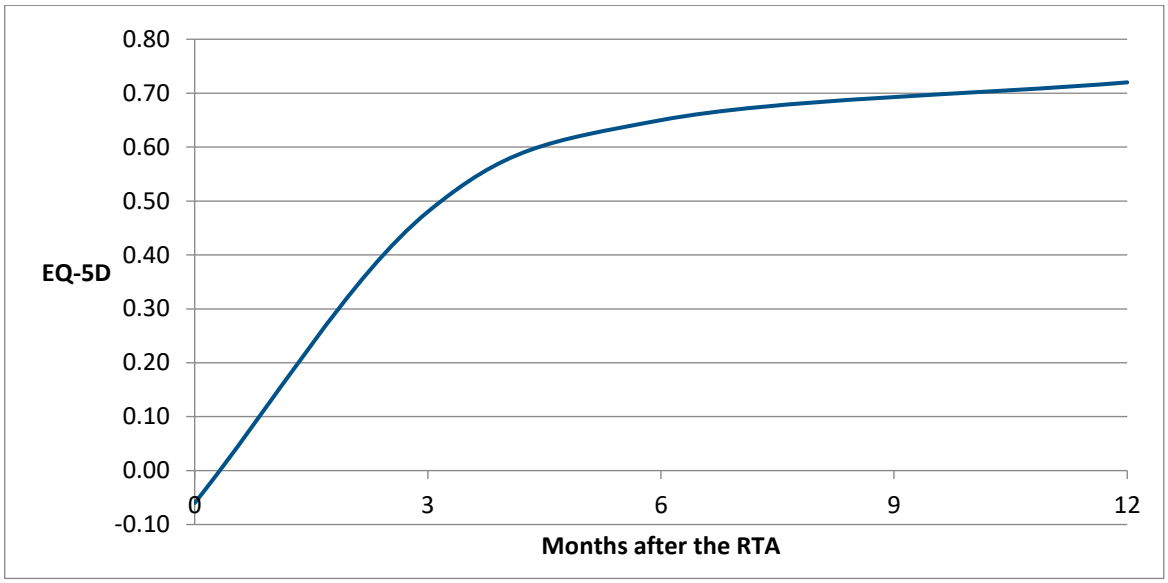
The utility loss caused by hip or post-hip fracture is calculated in each cycle looking at the number of people populating these two health states and subtracted from the overall QALYs calculated in that cycle.

2.3.6.4 Utility decrements after an RTA

Losses of utility after a road traffic accident were informed from a quality-of-life study[7] on a sample of patients being hospitalized after a road traffic injury in the UK. The study estimated a QALY loss for this sample equal to 0.21. Confidence intervals were not provided so for the probabilistic analysis we assumed that the standard error was equal to 20% of the mean.

This utility decrement was applied to people experiencing a severe road traffic accident for one cycle only as the study suggests that, by 12 months, the person has mostly recovered, and the quality of life significantly increased (see **Figure 8**).

Figure 8: EQ-5D Utility Scores over a 1-year follow-up period after a serious RTA



Source: Barner 2006[7]

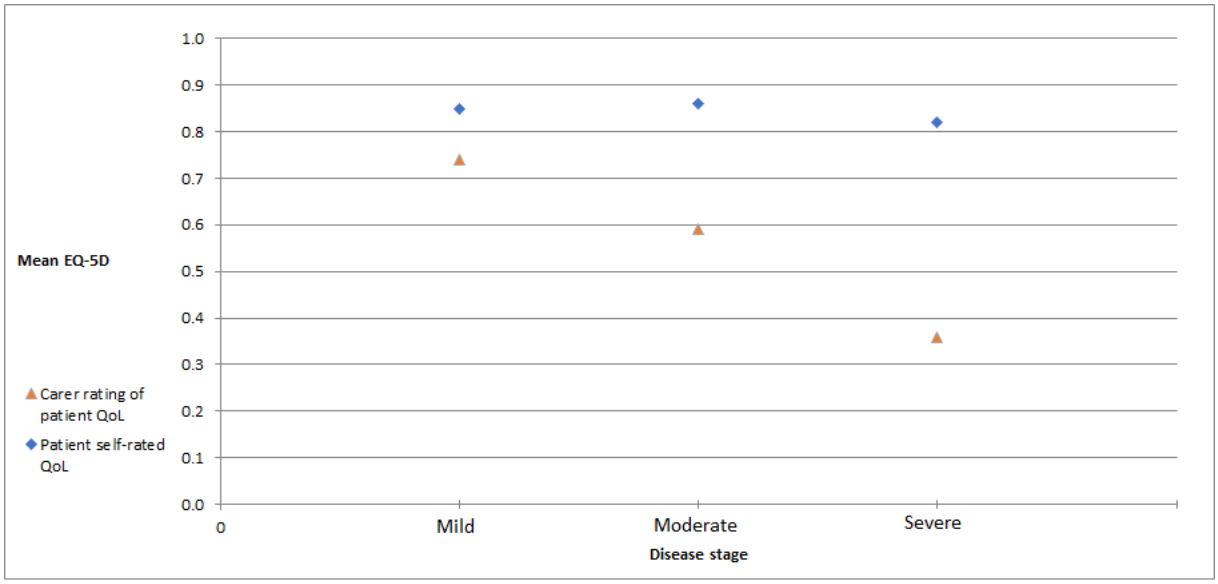
Fatal RTAs, of course, result in death and so a QOL decrement does not need to be defined. Slight road traffic accidents include a variety of injuries that often do not require medical treatment or only roadside attention[17]. Therefore, it was assumed that a slight RTA does not cause any QALY detriments in the short or long-term alike but only a cost (see section 2.3.7.6).

2.3.6.5 Utility of people with dementia

People in the dementia state have a lower quality of life due to the cognitive impairment caused by their condition. Data on their quality of life were taken from a literature review of 61 studies and measured in terms of EQ-5D 5L.[27] The studies were conducted in several countries, with only 15 in the UK, so not all used the UK tariff to measure EQ-5D. Nevertheless, the study was included in the model as it was assessed to be the most comprehensive assessment of the quality of life of people with dementia.

As quality of life is hard to self-assess for a patient affected by dementia, numerous studies reported proxy-rated utilities instead that were often collected from the person's main caregiver. **Figure 8** shows the mean EQ-5D reported both from the patients and carers for different levels of severity of dementia.

Figure 9: HR-QoL measured using the EQ-5D. Self- and proxy ratings in people with MCI or dementia by disease severity



Source: Landeiro 2020[27]

The figure shows that, as severity increases, the difference between self-rated health-related quality of life and carers' ratings of patient health-related quality of life becomes larger. Moreover, self-reported quality of life hardly varies across the different disease stages, which is rather unrealistic as severe forms of dementia are expected to significantly impact people and their ability to conduct daily life activities. By contrast, carers' ratings of patients' QoL show that quality of life decreases with the severity of the disease suggesting that severely impaired cognition may alter the ability to complete the questionnaire appropriately. Although NICE reference case recommends using measurement obtained patients, the committee decided to use caregivers proxy EQ-5D score in the base case scenario, although a sensitivity analysis using self-rated EQ-5D was included (see section 3.2). Table 17 shows EQ-5D rated by caregivers and patients for each level of severity of dementia.

Table 17: EQ-5D by dementia severity level (proxy and self-rated)

Severity of dementia	EQ-5D Caregiver as proxy [a]	EQ-5D Self-rated
Mild	0.74 (0.69 to 0.79)	0.85 (0.8 to 0.89)
Moderate	0.59 (0.47 to 0.71)	0.86 (0.76 to 0.96)
Severe	0.36 (0.18 to 0.53)	0.82 (0.64 to 1)

(a) These values were used in the base case scenario

Prevalence by age of the three severity levels of dementia, was informed from a recent cost study conducted in England[54] and is presented in Table 18.

Table 18: Severity of dementia by age

Age category	Mild dementia (%)	Moderate dementia (%)	Severe dementia (%)	Source
65-74	17.92%	40.83%	41.25%	Wittenberg 2019[54]
75-84	18.97%	40.18%	40.84%	
85+	14.70%	31.53%	53.78%	

At each cycle, the model applies the EQ-5D from Table 17 to people in the dementia 1st year or dementia long-term states using the prevalence reported in Table 18. Utility values are not assigned directly but firstly compared with UK reference utilities reported by Ara and Brazier[3]

using as reference age the mean age reported in the meta-analysis. The comparison allowed us to calculate a utility multiplier, which was applied to the Ara and Brazier utility vector to calculate an age-specific utility vector for people with dementia.

2.3.7 Resource use and costs

2.3.7.1 Intervention

The cost of CBT + TO and TOA interventions were calculated by looking at the description of the trials included in the clinical review. As the estimation of the cost varies consistently across the three different trails, an average of the costs of the three was used in the base case scenario whereas each different estimation was used in the scenario sensitivity analyses. Table 19 shows the data used to calculate the cost of CBT and TO.

Table 19: Cost of CBT and TO interventions

	Baillargeon [4]	Morin [33]	Oude Voshaar [47]	Base case scenario (average)
Group cognitive behavioural therapy without tapering				
Group size	6	5	4	5
Number of sessions	9	10	5	8
Duration (hour)	1.5	1.5	2.0	1.7
Cost per hour [a]	£61	£61	£61	£61
CBT cost (per patient)	£137	£183	£153	£158
Tapering off				
Group size	1	1	1	1
Number of visits to GP	8	10	4	7.33
Duration (hour)	0.25	0.25	0.25	0.25
Cost per hour [b]	£153	£153	£153	£153
TO cost (per patient)	£306	£383	£153	£281

(a) PSSRU[12] Band 7 cost including qualification costs
(b) PSSRU[12] GP hourly cost including direct care and qualification costs

The total cost of each intervention used in the base case scenario of the model is illustrated in Table 20. Naturally, the cost of CBT+TO was assumed to be equal to the sum of the cost of the two strategies whereas the cost of usual care was assumed to be equal to £0.

Table 20: Cost of each strategy (base case scenario)

Strategy	Cost	Source
CBT+TO	£438	Bailargeon 2003[4]
TOA	£281	Morin 2004[33]
Usual care	£0	Oude Voshaar 2006[47]

2.3.7.2 Benzodiazepines

The cost of each benzodiazepine was obtained from the British National Formulary (BNF).[8] A list of all the benzodiazepines included in the model is presented in Table 21. Dosages were

taken from BNF whereas the average cost per mg was calculated using the Prescription Cost Analysis 2019 database.[42]

Table 21: Benzodiazepines - cost

Benzodiazepines	Daily dosage(in mg)[a]	Cost per mg[b]	Cost per day	Weight[c]
Loprazolam	1.5	£0.80	£1.21	0.02
Lormetazepam	1.5	£0.71	£1.07	0.01
Nitrazepam	7.5	£0.02	£0.18	0.13
Temazepam	15	£0.005	£0.07	0.31
Diazepam	22.5	£0.01	£0.19	0.33
Chlordiazepoxide	30	£0.02	£0.70	0.01
Lorazepam	2.5	£0.12	£0.30	0.19
Oxazepam	90	£0.02	£1.74	0.01

(a) Source: BNF[8]
(b) Source: PCA - drug tariff[42]
(c) Calculated looking at "days of dosage" sold in PCA dataset

To obtain the daily cost of the overall benzodiazepine class, the Prescription Cost Analysis database was used to convert in "days of dosage", the quantity of benzodiazepines sold in England each year. Days of dosage were thus used as weights to calculate the weighted average cost of benzodiazepines. This gave us a daily cost of £0.20 and an annual cost of £77.84.

2.3.7.3 Cost of falls

The cost of a fall in England was calculated using a Scottish cost analysis from Craig 2013[11]. Data on the number of people who fall attending a GP, requiring ambulance service, A&E, hospitalisation and subsequent care home residence were obtained from the Information Services Division (ISD) and Scottish Ambulance Service (SAS). Table 22 illustrates all fall-related events. 20% of people who fall were estimated to experience a "serious fall" by summing those attending general practices, calling an ambulance and attending A&E.

Table 22: Falls-related events proportion

Event	Proportion	Source
Serious fall	20% of total falls	Craig 2013[11]
GP attendance	51% of serious falls	
Ambulance call-out	61% of serious falls	
A&E attendance	80% of serious falls	
Admissions	35% of A&E attendances	
Re-admissions	7% of admissions	
Discharged at home	64% of admissions	
Discharged at residential: short-term	21% of admissions	
Discharged at residential: long-term	15% of admissions	

The cost of each event was estimated using standard UK sources (NHS Reference Costs and PSSRU) and the study from Craig. The cost of a GP visit was informed from PSSRU. The cost of an ambulance call-out was obtained from NHS Reference Costs 2018-2019. The cost of A&E attendance was calculated separately for those who are admitted and for those

who are not admitted by taking an average, weighted by the number of attendances of all the items listed in the NHS Reference Costs under A&E admitted and non-admitted. Finally, the costs of an inpatient stay, home discharge, or discharge to a residential care centre were informed from Craig 2013[11] and inflated to current prices excluding the costs associated with hip fracture, as these were reported separately in the study. All the costs used in the model and their sources are listed in Table 23.

Table 23: Falls-related events cost

Event	Cost	Source
GP visit [a]	£38	PSSRU 2020[12]
Ambulance call-out	£257	NHS Ref Costs 2018-2019[43]
A&E non admitted	£84	NHS Ref Costs 2018-2019[43]
A&E Admitted	£195	NHS Ref Costs 2018-2019[43]
Inpatient stay (no HF)	£8,018	Craig 2013[11] inflated to 2018/19 prices
Home discharge	£1,923	Craig 2013[11] inflated to 2018/19 prices
Residential: Short-term	£9,101	Craig 2013 inflated to 2018/19 prices
Residential: long-term	£71,393	Craig 2013 inflated to 2018/19 prices

(a) Including direct care and qualification costs

Data presented in Table 22 and Table 23 allowed us to estimate the average cost of a fall in the UK which amounts to £1,294.

2.3.7.4 Cost of hip fracture

The cost of a hip fracture in England was estimated using a UK cost analysis[28] conducted on a cohort of 33,152 patients admitted with a hip fracture in a UK region between 2003 and 2013 and identified from hospital records and followed until death or administrative censoring.

The analysis estimated the cost occurring in the first year after a hip fracture and the cost occurring in the second year. They are both presented in Table 24. Costs included diagnostic and treatment cost both outpatient and inpatient sustained by the NHS.

Table 24: Hip fracture cost

Hip fracture cost	Mean	Standard error	Source
First year	£14,971	£254	Leal 2016[28] inflated to 2018/19 prices
Second year	£2,260	£90	

The model applies the first-year cost to those in the hip fracture tunnel state and the second-year cost to those in the post-hip fracture state. This naturally assumes that the same cost sustained in the second year will be sustained in the following years as well. This hypothesis appears to be supported by the original study[28] which found that hospitalisation costs remain stably above pre hip fracture cost until the last follow-up (24 months) without showing any downward trend. Although this does not prove that those costs will be sustained lifelong, it is likely that some hospitalisation costs will be sustained in the long-term as hip fractures have been shown to affect the life of people for several years.[16, 23]

2.3.7.5 Cost of dementia

The cost of dementia was estimated using an English cost study from 2015[54]. The study measured the average annual cost of people with dementia derived from multiple UK sources during the first and the second year after the diagnosis and grouped in three categories of severity of dementia: mild, moderate and severe. For the purpose of this

analysis, only the costs relevant to the NHS were extracted. Those were the primary and secondary care costs, as well as the proportion of social care cost sustained by NHS, which is equal, according to the authors, to the 39.4% of the overall social care costs. The costs were inflated to 2018/19 prices and are shown in Table 25.

Table 25: Cost of dementia by severity and year [a]

Year after the diagnosis	Mild	Moderate	Severe
First year	£3,478	£4,457	£7,247
Second year	£4,053	£6,026	£7,844

(a) 2015 prices inflated to 2018/19

The model applies the costs presented in table 25 using the prevalence data from Table 18 to estimate the number of people with mild, moderate and severe dementia. It was assumed that the cost sustained in the second year will be sustained for the rest of the life of people with dementia. This appears to be plausible as symptoms of dementia are not expected to decrease over time and they may even increase if the condition gets worse.

2.3.7.6 Cost of RTAs

The Department for Transport have 2018 data on the cost of RTAs from a healthcare perspective (Medical and Ambulance) disaggregated according to the severity of the casualty.[19] These are reported in Table 26.

Table 26: Medical and ambulance cost per road traffic accident casualty

Injury Type	Total Casualties	Total Costs	Cost per casualty (£)
Slight Injury	133,302	£146m	£1,095
Seriously Injured	25,511	£401m	£15,719
Fatal	1,784	£11m	£6,166

In the sensitivity analysis, the cost of the police was added to the healthcare cost of RTAs (see section 2.5.8).

2.4 Computations

The model was constructed in Microsoft Excel 2010 and was evaluated by cohort simulation. Time dependency was built in by cross-referencing the cohort's age as a risk factor for mortality. Quality of life scores were adjusted for age to reflect the declining trend of quality of life over time. Utility in CBT + TO, TOA and usual care were assumed to converge to the same value (usual care) 2 years after the intervention in the base case scenario.

People started in the Markov model in the CBT + TO, TOA or Usual Care arms either as abstinent or under benzodiazepines. People then moved to the other health states (hip fracture or dementia) based on probabilities of events occurring which was calculated using incidence data and benzodiazepine relative effects. Mortality transition probabilities in the Markov model depend on the health states people are in. Each cycle, instant events (falls and road traffic accidents) are calculated based on the population at risk and used to calculate loss of QALYs and costs.

Mortality rates were converted into transition probabilities for the respective cycle length (1 year in the base case) before inputting into the Markov model.

$Transition\ Probability\ (P) = 1 - e^{-rt}$	Where *r*=selected rate *t*=cycle length (1 year)

To calculate QALYs for each cycle, life years were weighted by a utility value which was treatment dependent. A half-cycle correction was applied, assuming that people transitioned between states on average halfway through a cycle. QALYs were then discounted at 3.5% to reflect time preference. QALYs during the first cycle were not discounted. The total discounted QALYs were the sum of the discounted QALYs per cycle.

Costs per cycle were calculated on the same basis as QALYs and were discounted at 3.5% to reflect time preference. Each of the health states had specific costs applied.

Discounting formula:

$$\text{Discounted total} = \frac{\text{Total}}{(1+r)^n}$$

Where:
r=discount rate per annum
n=time (years)

In the deterministic and probabilistic analyses, the total cost and QALYs accrued by each cohort was divided by the number of patients in the population to calculate a cost per patient and cost per QALY.

2.5 Sensitivity analyses

In addition to the probabilistic sensitivity analysis, a range of one-way sensitivity analyses were undertaken. These are shown in table 27, where the scenarios used in the base case scenario are highlighted in green:

Table 27: Scenario analyses

Feature	Scenarios	Description
Age	50 years old	Initial age is 50
	64 years old	Initial age is 64 (average of mean age from the trials)
	70 years old	Initial age is 70
Relapse rate	Same relapse rate in the first and second year	The relapse rate is calculated using data from the trials post-intervention and at 1-year follow-up
	Relapse rate decreasing in the second year	Relapse in the first year is calculated using trials data. Relapse in the second year is calculated using the survival analysis from Morin [34]
Relapse in usual care	No relapse in usual care	Relapse is allowed only in the CBT+TOA and TOA arms
	Relapse in usual care	Relapse rate in the UC arm is assumed to be equal to relapse in TOA

Relapse duration	Relapses for 2 years	Relapses occur during the first two years. Beyond the second year, relapses are assumed not to occur anymore
	Relapse for 5 years	Relapses occur during the first five years. Beyond the fifth year, relapses are assumed not to occur anymore
Dose reduction	Savings from dosage reduction not included	Additional savings due to dose reduction in those under benzodiazepine are not included
	Savings from dosage reduction included	The proportion of people achieving a dose reduction described in Baillargeon 2003[4] is used to calculate pharmaceutical saving
RTA OR	Same OR applied to all age groups	Overall OR calculated from Barbone[5] applied to all age groups
	Age-specific OR	Age-specific OR calculated from Barbone[5] applied according to age
Suicide	Benzodiazepine does not affect suicide	Benzodiazepines are assumed not to affect suicide risk
	Benzodiazepine affects suicide	Benzodiazepine users are at a higher risk of suicide
Utility convergence	No convergence in utility	Differences in utility scores found after, at 18 months, are assumed to last for the rest of the life
	Convergence after 1 year	Differences in utility scores found after at 18 months converge in the second year
	Convergence after 5 years	Differences in utility scores found after, at 18 months, converge in the sixth year
EQ-5D in people with dementia	EQ-5D reported by caregivers	EQ-5D scores in people with dementia are reported by caregivers
	EQ-5D self-reported	EQ-5D scores in people with dementia are reported by the people themselves

Intervention costs	Average of the trials	Cost of the intervention calculated using the average cost from the trials
	Baillargeon cost	Cost of the intervention based on Baillargeon 2003[4]
	Morin cost	Cost of the intervention based on Morin 2004[33]
	Oude Voshaar cost	Cost of the intervention based on Oude Voshaar 2006[47]
Police cost	Police cost not included in RTA cost	Calculation of RTA costs included only health care costs and not police costs
	Police cost included in RTA cost	Calculation of RTA costs included health care costs as well as police costs

In the next sections, all the scenario analyses and their rationale are explained

2.5.1 Age

There was some uncertainty around the starting age of the cohort. The average age from the trials was used in the base case scenario, although it was noted that it may be higher than the average age of people taking benzodiazepines in the UK. No study reporting the average age of benzodiazepine users was identified in the UK, but a Spanish study found an average age of 56. If prescription patterns in the UK are similar, we may expect that the average age of a benzodiazepine user in the UK is between 50 to 60 as well. Therefore, a scenario analysis where the starting age of the cohort is 50 was added. In addition, a scenario analysis with a starting age of 70 was used to test the results of the model in an older cohort.

2.5.2 Relapse

In the base case scenario of the model, the relapse rate in the second year was calculated using the survival analysis on relapse made by Morin.[34] However, as the survival analysis had a relatively small sample size, an alternative scenario was tested where the relapse rate in the second year is the same rate used for the first year, which was calculated through the clinical meta-analysis.

Relapse in the usual care group was assumed to be zero in the base scenario, as the only trial comparing CBT and TO with usual care did not find any relapse occurring in this group[47]. However, it was noted that the sample size in the trial was small and that, in the real world, some who discontinued benzodiazepines by themselves, may relapse as well. Therefore, a scenario analysis was included where the relapse rate in usual care was assumed to be equal to the relapse rate in people in TOA arm.

Finally, as mentioned before, relapse into drug use was allowed for 2 years as recommended by published evidence and committee's experience. However, an alternative scenario where relapse can occur for 5 years was tested, using as yearly probability the average of the relapse probabilities during the first two years. This was done to assess the impact of this assumption on the results of the model.

2.5.3 Dose reduction

One trial included in the review[5] showed that people not achieving cessation were, in most cases, still able to reduce consistently their daily dosage of benzodiazepines. The quality of the data was considered too weak for this to be included in the main analysis, but a scenario analysis was added where data on people achieving a reduction >50% from Baillargeon 2003[4] were used to calculate savings due to the effect of the treatment in reducing the dosage of benzodiazepines.

2.5.4 Road traffic accident (RTA) odds ratio

In the base case scenario, the overall odds ratio found in the study from Barbone 1998[5] was applied to all age groups. In the sensitivity analysis, two different odds ratios were tested for two different age groups extracted by the same evidence:

- 50-64: 1.48
- >65: 1

2.5.5 Suicide

In one of the sensitivity analyses, benzodiazepines were assumed to increase suicides. However, similarly with RTA, (see 2.3.5.4), the overall probability of committing suicides could not be added to the model as deaths due to suicide are already captured by the life tables. Instead, as in the case of fatal RTA, the excess suicide mortality caused by benzodiazepines had to be added to the mortality in people in the benzodiazepine state. The excess mortality was calculated by subtracting the probability of suicides occurring in the general population from the probability of suicide occurring in people taking benzodiazepines. This latter was calculated by applying the odds ratio shown in Table 8 to the odds of dying because of suicide (which was then converted back to a probability). The resulting probability, or excess mortality, is the additional probability of dying because of suicide caused by benzodiazepines.

2.5.6 Utility convergence

There was some uncertainty regarding the utility score in the three arms. In particular, the difference in the utility in the CBT, TOA and UC arms at 18 months, which was considered to be due to long-term withdrawal symptoms in the two intervention arms, was expected to decrease over time. In the base case scenario it was assumed that, by year 2, people in each group would have the same quality of life score (see 2.3.6.1). In the sensitivity analysis two different scenarios were tested:

- Utility scores converge 5 years after the intervention
- Utility scores do not converge

These scenarios were tested to assess the impact of the assumption on quality of life in the base case scenario on the model's results. The committee thought that it is very unlikely that any quality-of-life harm caused by withdrawal symptoms, or the intervention would last for a period longer than 2 years, therefore these two assumptions were used only in the sensitivity analysis.

2.5.7 EQ-5D in people with dementia

In the base case scenario, mean EQ-5D in people with dementia was calculated using proxy ratings collected from main caregivers. The NICE reference case recommends using people self-reported quality of life measures when possible and using proxy values only in the absence of self-rated data. However, a disease like dementia, especially in the most

advanced stages, may affect the ability of people to complete the questionnaire. This may have happened in the studies included in the systematic review used as evidence[27] as, surprisingly, self-reported EQ-5D for mild, moderate and severe dementia are relatively high and similar. This was considered to be rather unrealistic by the committee as people experiencing severe forms of dementia often struggle to carry out daily tasks without the help of a caregiver. Therefore, it was decided to use proxy EQ-5D in the base case scenario and to test self-rated values only in a sensitivity analysis.

2.5.8 Intervention cost

The costs of the two interventions in the base case scenario are the average costs across the three trials included in the clinical review[4, 33, 47] that were calculated looking at the description of health care resource use reported by each trial. Three scenarios were tested where the costs of the interventions were calculated using each study separately. These costs are the following:

- Baillargeon 2003[4]
  - CBT+TO = £443
  - TOA = £306
- Morin 2004[33]
  - CBT+TO = £556
  - TOA = £383
- Oude Voshaar 2006[47]
  - CBT+TO = £306
  - TOA = £153

2.5.9 Police cost included in the RTA cost

In the base case scenario, only healthcare costs derived by a road traffic accident were incorporated in the model. The Department of Transport included other costs as well, such as costs related to police intervention and investigation[19]. To include a broader societal perspective, in the sensitivity analysis police costs were added to healthcare RTA costs.

2.6 Model validation

The model was developed in consultation with the committee; model structure, inputs and results were presented to and discussed with the committee for clinical validation and interpretation.

The model was systematically checked by the health economist undertaking the analysis; this included inputting null and extreme values and checking that results were plausible given inputs. The model was peer-reviewed by a second experienced health economist from the NGC; this included systematically checking many of the model calculations.

2.7 Estimation of cost effectiveness

The widely used cost-effectiveness metric is the incremental cost-effectiveness ratio (ICER). This is calculated by dividing the difference in costs associated with 2 alternatives by the difference in QALYs. The decision rule then applied is that if the ICER falls below a given cost per QALY threshold, the result is considered to be cost-effective. If both costs are lower and QALYs are higher, the option is said to dominate and an ICER is not calculated.

$$ICER = \frac{Costs(B) - Costs(A)}{QALYs(B) - QALYs(A)}$$

Cost effective if:
- ICER < Threshold

Where: Costs(A) = total costs for option A; QALYs(A) = total QALYs for option A	

When there are more than 2 comparators, as there is in this analysis, options must be ranked in order of increasing cost then options ruled out by dominance or extended dominance before calculating ICERs excluding these options. An option is said to be dominated, and ruled out if another intervention is less costly and more effective. An option is said to be extendedly dominated if a combination of 2 other options would prove to be less costly and more effective.

It is also possible, for a particular cost-effectiveness threshold, to re-express cost-effectiveness results in terms of net monetary benefit (NMB). This is calculated by multiplying the total QALYs for a comparator by the threshold cost per QALY value (for example, £20,000) and then subtracting the total costs (formula below). The decision rule then applied is that the comparator with the highest NMB is the cost-effective option at the specified threshold. That is the option that provides the highest number of QALYs at an acceptable cost.

$Net\,Monetary\,Benefit(X) = (QALYs(X) \times \lambda) - Costs(X)$ *Where: λ = threshold (£20,000 per QALY gained)*	Cost effective if: • Highest net benefit
$Net\,Health\,Benefit(X) = (QALYs(X)) - Costs(X)/\lambda$ *Where: λ = threshold (£20,000 per QALY gained)*	Cost effective if: • Highest net benefit

Both methods of determining cost -effectiveness will identify exactly the same optimal strategy. For ease of computation, NMB is used in this analysis to identify the optimal strategy.

Results are also presented graphically where total costs and total QALYs for each diagnostic strategy are shown. Comparisons not ruled out by dominance or extended dominance are joined by a line on the graph where the slope represents the incremental cost-effectiveness ratio.

2.8 Interpreting results

NICE sets out the principles that committees should consider when judging whether an intervention offers good value for money.[36-38] In general, an intervention was considered to be cost-effective if either of the following criteria applied (given that the estimate was considered plausible):

- The intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more clinically effective compared with all the other relevant alternative strategies), or
- The intervention costs less than £20,000 per quality-adjusted life-year (QALY) gained compared with the next best strategy.

As we have several interventions, we use the net monetary benefit (NMB) to rank the strategies on the basis of their relative cost-effectiveness. The highest NMB identifies the optimal strategy at a willingness to pay of £20,000 per QALY gained.

3 Results

3.1 Base case

The probabilistic results of the base case scenario are presented in Table 28, which provides a breakdown of costs per patient.

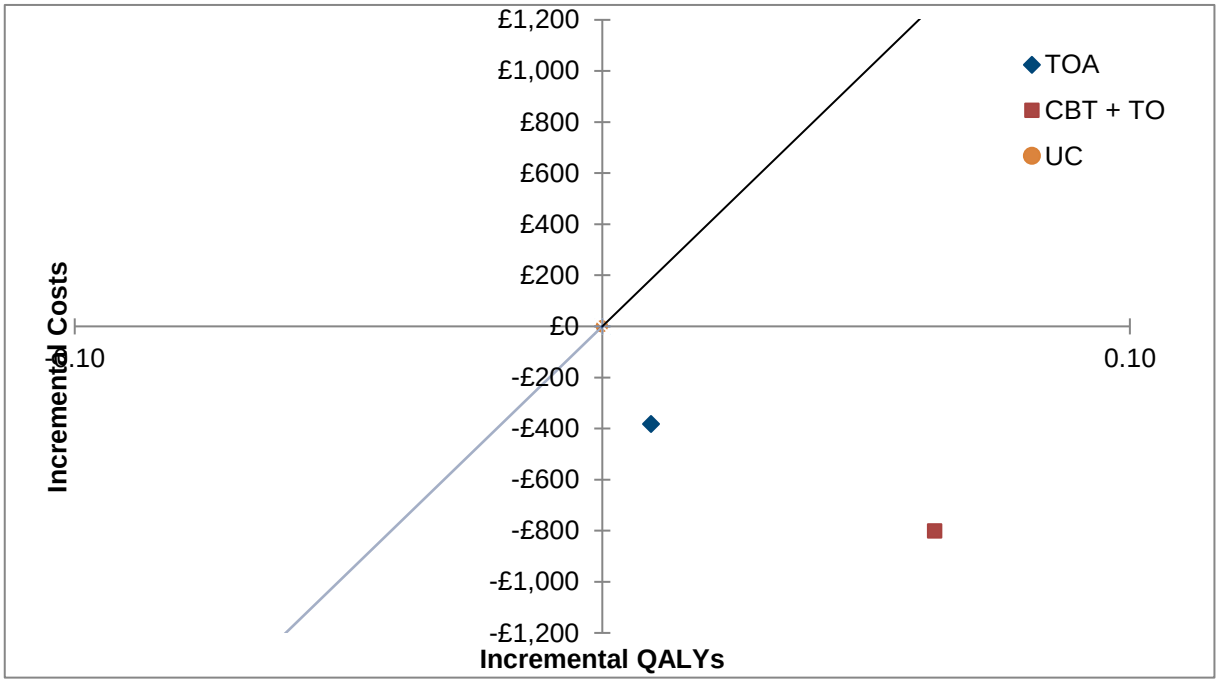
Table 28: Base case scenario probabilistic results[a]

	CBT + TO	TOA	Usual Care
Intervention cost	£438	£281	£0
Benzodiazepine cost	£473	£613	£775
Fall Injuries cost	£8,177	£8,459	£8,781
Hip fractures cost	£1,500	£1,503	£1,505
RTA cost	£58	£61	£66
Dementia cost	£3,807	£3,955	£4,127
Total cost	£14,453	£14,872	£15,254
QALYs	8.63	8.57	8.57
NMB (20k)	**£158,113**	**£156,620**	**£156,053**
INMB rank	**1**	**2**	**3**

1. *Costs and QALYs are mean values per patient*

INMB ranking indicates that CBT+TO is the most preferred strategy, followed by TOA and then usual care. Both CBT+TO and TOA dominate usual care, meaning that they are associated with a lower cost and a higher quality of life, as shown in the cost-effectiveness plane in Figure 10.

Figure 10: Cost-effectiveness plane



The probability that CBT + TO is cost-effective compared to usual care is 86% at a £20,000 threshold and 87% at a £30,000 threshold. The probability that TOA is cost-effective compared to usual care is 62% at a £20,000 threshold and 65% at a £30,000 threshold.

Finally, CBT+TO is cost-effective against TOA in 80% of the simulations at a £20,000 and £30,000 thresholds alike.

3.2 Scenario analysis

Several one-way sensitivity analyses were conducted as mentioned in section 2.5. The deterministic results of each scenario are illustrated in Table 29.

Table 29: Deterministic results of the scenario analysis

Scenarios	CBT vs TOA INMB[a]	CBT vs UC INMB[a]	TOA vs UC INMB[a]	Ranking
Base case	£1,473	£2,209	£737	1. CBT + TO 2. TOA 3. Usual care
Starting age = 50	£1,236	£1,538	£302	1. CBT + TO 2. TOA 3. Usual care
Starting age = 70	£1,389	£2,075	£686	1. CBT + TO 2. TOA 3. Usual care
Same relapse factor in the second year	£701	£1,179	£479	1. CBT + TO 2. TOA 3. Usual care
5 years relapse duration	£1,047	£698	-£348	1. CBT + TO 2. Usual care 3. TOA
Relapse in UC equal to relapse in TOA	£1,473	£2,597	£1,124	1. CBT + TO 2. TOA 3. Usual care
Savings due to dose reduction included	£1,438	£2,289	£852	1. CBT + TO 2. TOA 3. Usual care
Age-specific OR for RTA	£1,473	£2,209	£737	1. CBT + TO 2. TOA 3. Usual care
Benzodiazepine increases suicide risk	£1,473	£2,209	£737	1. CBT + TO 2. TOA 3. Usual care
No convergence in utility	£6,059	£730	-£5,329	1. CBT + TO 2. Usual care 3. TOA
Convergence after 5 years	£2,941	£1,739	-£1,203	1. CBT + TO 2. Usual care 3. TOA
Self-reported EQ-5D for dementia	£1,318	£1,852	£534	1. CBT + TO 2. TOA 3. Usual care
Baillargeon intervention cost	£1,493	£2,204	£711	1. CBT + TO 2. TOA 3. Usual care

Scenarios	CBT vs TOA INMB[a]	CBT vs UC INMB[a]	TOA vs UC INMB[a]	Ranking
Morin intervention cost	£1,447	£2,082	£635	1. CBT + TO 2. TOA 3. Usual care
Oude Voshaar intervention cost	£1,478	£2,342	£864	1. CBT + TO 2. TOA 3. Usual care
Police cost included in RTA cost	£1,474	£2,212	£738	1. CBT + TO 2. TOA 3. Usual care

(a) INMB calculated using a threshold of £20,000

CBT was found to be the most preferred strategy in all scenarios. Only in the two scenarios where utilities converge after 5 years, or do not converge at all, and in the scenario with relapses lasting for 5 years the ranking changed, with usual care becoming cost-effective compared with TOA, although CBT remained the most effective strategy. This suggests that even in the scenario where the loss of utility due to withdrawal symptoms lasts for a long time period, CBT is effective in reducing those symptoms and ultimately making the intervention cost-effective.

4 Discussion

4.1 Summary of results

One original cost-utility analysis found that group CBT + tapering off dominates both tapering off alone and usual care. The analysis was assessed as directly applicable with minor limitations.

4.2 Limitations and interpretation

The analysis demonstrated that group cognitive behavioural therapy alongside tapering off is cost-effective compared to tapering off alone and usual care. The strategy was found to dominate the other two strategies, as it is associated with a lower healthcare cost and higher quality of life. The results are robust to all the assumptions used in the model as in none of the scenarios tested CBT + TO was found to be less cost-effective than the alternatives.

CBT + TO was found to cost less than the other strategies in the long-term even though this strategy has the highest initial cost. Significant downstream savings were found for benzodiazepine consumption, fall injury treatment cost and dementia care cost. In addition, as CBT + TO was found to reduce cases of fall injuries, hip fracture, road traffic accidents and dementia, the strategy was associated with a slightly higher overall quality of life.

This study is subject to some limitations. Firstly, it was assumed that health states do not overlap meaning that a person with hip fracture cannot develop dementia and vice versa. This was made for practical purposes although, in the real world, it is obvious that these two conditions co-exist. Furthermore, it is possible that people with dementia have a higher risk of falling and, consequently, incur in a hip fracture. Although this may indeed underestimate the number of falls or hip fractures predicted by the model, people with dementia are associated with long-term hospital costs and loss of utility collected from real-world evidence, which implies that costs and quality of life decrements caused by other consequences of dementia (including falls and fractures) are already captured in the model.

Secondly, it was assumed that people moving to dementia or hip fracture health states would immediately withdraw from benzodiazepines. This was done for modelling purposes as well as to reflect the fact that prescribers usually avoid giving psychoactive drugs to people with poor mobility or dementia. The NICE guideline on dementia[35] recommends using non-pharmacological methods to manage non-cognitive symptoms in dementia, and to start a low dose of antipsychotics only if the other treatments fail and if the person is at risk of harming themselves or others or experiencing hallucinations. However, in practice, benzodiazepines are often prescribed even to people with low mobility or cognitive disorders. A Swedish study[26] found, for instance, that people admitted to the hospital for a hip fracture are more likely to receive a prescription of antipsychotic (including benzodiazepine) 6 months after the event. Although people with dementia or hip fracture represent a relatively small population in the model, it is possible that, particularly later in life, the model is underestimating the real number of people with dementia or hip fracture on benzodiazepines. This implies that CBT + TO and TOA may be even more cost-effective than usual care, as they prevent cases of dementia and hip fracture and, consequently, avoid additional benzodiazepine prescriptions.

Thirdly, dementia and hip fracture were modelled as permanent states, meaning that a person cannot recover after experiencing one of these two states. This is justifiable for dementia as, unfortunately, dementia cannot be cured, and it is expected to worsen over time. On the other hand, it is possible that a person with a hip fracture will fully recover after a certain period of time. Evidence used for the model suggest that mortality in people with a hip fracture is higher than the general population at least 10 years after the event[23]. The hazard ratio stays consistently over 1 for the entire follow-up period of the study suggesting

that mortality will remain higher even beyond the last follow-up. Moreover, evidence[31] show that hip fracture may trigger or accelerate frailty in old people, causing effects on health lasting, perhaps, for the rest of their life. Nevertheless, it is possible that the model is overestimating the impact of hip fracture for those who fully recover, although this should not affect the results of the model in a meaningful way as the number of hip fractures predicted by the model is relatively low.

Utility scores were taken from Oude Voshaar 2006[47] who collected mean utility score in each arm without distinguishing between people who discontinued or remained on benzodiazepine. Consequently, we were not able to assign two specific utility values to people in the abstinent or “on benzodiazepine” states. This would have been, naturally, desirable as the committee highlighted the importance of capturing losses of utility caused by withdrawal symptoms in those who are abstinent. Instead, the mean value was assigned to each arm regardless of the number of people in the abstinent or “on benzodiazepine” state. This should still ensure that the lower utility in the abstinent states is captured as we applied the mean utility from the study, which had abstinent and people under benzodiazepine in different proportions in each arm. However, issues may have arisen during the probabilistic sensitivity analysis as, even though the proportions of people in each state varied, the mean utility scores did not vary accordingly. Moreover, Oude Voshaar[47] was the only trial failing to find CBT effective compared with TOA in increasing the cessation rate. This implies that the study may overestimate utility of people receiving CBT, as the committee expect more people to be abstinent in this group and, therefore, to be at risk of experiencing withdrawal symptoms.

A higher QoL score was observed in the usual care group at 18 months [47]. The committee expected people coming off from long-term benzodiazepine medication to have lower quality of life due to withdrawal symptoms, but it was surprising to see a difference so long after the intervention (even if this was based on changes in SF-36 that were not statistically significant). The duration of disutility caused by the withdrawal symptoms is uncertain and therefore three scenarios were analysed where the convergence of utility between the different strategies occurs at 2 and 5 years after the intervention or does not occur at all. In all three scenarios, CBT + TO remained cost effective compared to TOA and usual care; therefore, the conclusions of the model seem robust.

Finally, the model did not include early-onset of dementia (<65 years old), or death caused by falls for reasons other than hip fracture. The first assumption was justified as there was not sufficient data to model early cases of dementia. These are relatively rare events so should not impact the model significantly. Falls can cause deaths for reasons other than hip fracture, for instance, head injury due to a fall. However, these were not included in the model to avoid double-counting mortality, and due to the limited availability of data. It is possible therefore that the model is underestimating mortality in people experiencing a serious fall although, again, this is not expected to affect the model significantly.

4.3 Generalisability to other populations or settings

The analysis is based on people taking benzodiazepines (for at least 3 months) who were unable to stop by themselves. In some of the trials included, people were first asked to stop using benzodiazepines through a letter by their GP to ensure that the population enrolled truly needed additional services to help them discontinue the drug. This highlights the fact that CBT + TO should be offered only to those whose previous discontinuation attempts through GP consultation or letter failed. Offering CBT + TO to people who can discontinue without additional support would obviously not be cost-effective.

This analysis is based on a UK population using UK incidence data and costs, so it may not be transferable to other jurisdictions where costs and incidence data are different.

4.4 Comparisons with published studies

To our knowledge, the only health economics analysis evaluating CBT + TO was conducted by Oude Voshaar[48] and is a within-trial analysis based on one of the randomised controlled trials included in the clinical review.[47] The cost-effectiveness analysis found CBT + TO to be extendedly dominated by TOA. This is because the analysis was based on a single trial which did not find any evidence of effectiveness for CBT alongside TO. Conversely, the guideline model is based on a meta-analysis including all three trials available in the literature which found CBT to increase the number of people discontinuing benzodiazepine (see forest plots in appendix A).

A modelling analysis comparing benzodiazepine cessation strategies was conducted by Moriarty.[32] The analysis found that any treatment that makes people on benzodiazepine discontinue their medication with a 100% cessation rate would result in savings equal to £2,971 and a quality-of-life gain of 0.07 QALYs per person. This is fairly aligned with the model in terms of cost as, if we assume that CBT is 100% effective, the savings predicted by the model are £2,382. On the other hand, the model predicts a higher gain in terms of QALYs, 0.25, compared to the 0.07 of Moriarty 2019. This is most likely due to the high negative impact on quality of life associated with dementia, which is an outcome in the guideline model but not in the Markov model of Moriarty.[32]

4.5 Conclusions

This economic evaluation demonstrated that, group CBT alongside tapered withdrawal dominates tapering withdrawal alone and usual care, in people using continuous benzodiazepine medication and who are not able to discontinue themselves.

The conclusions of the analysis are robust to all the assumptions used. To our knowledge, this is the first long-term health economic analysis evaluating CBT + TO for people using benzodiazepines. Previous research finding CBT not cost-effective, was based on a single trial and did not look at the long-term consequences and adverse events related to benzodiazepine use.

References

1. Age UK. Falls in later life: a huge concern for older people. 2019. Available from: https://www.ageuk.org.uk/latest-press/articles/2019/may/falls-in-later-life-a-huge-concern-for-older-people/ Last accessed: 06 August 2021.

2. Ara R, Brazier J. Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). Value in Health. 2008; 11(7):1131-1143

3. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value in Health. 2010; 13(5):509-518

4. Baillargeon L, Landreville P, Verreault R, Beauchemin JP, Gregoire JP, Morin CM. Discontinuation of benzodiazepines among older insomniac adults treated with cognitive-behavioural therapy combined with gradual tapering: a randomized trial. CMAJ: Canadian Medical Association Journal. 2003; 169(10):1015-1020

5. Barbone F, McMahon AD, Davey PG, Morris AD, Reid IC, McDevitt DG et al. Association of road-traffic accidents with benzodiazepine use. The Lancet. 1998; 352(9137):1331-1336

6. Barendregt JJ. The effect size in uncertainty analysis. Value in Health. 2010; 13(4):388-391

7. Barnes J, Thomas P. Quality of life outcomes in a hospitalized sample of road users involved in crashes. Annual proceedings Association for the Advancement of Automotive Medicine. 2006; 50:253-268

8. BMJ Group and the Royal Pharmaceutical Society of Great Britain. British National Formulary. Available from: https://www.evidence.nhs.uk/formulary/bnf/current Last accessed: 30/06/2021.

9. Cato V, Holländare F, Nordenskjöld A, Sellin T. Association between benzodiazepines and suicide risk: a matched case-control study. BMC Psychiatry. 2019; 19(1):317

10. Church J, Goodall S, Norman R, Haas M. The cost-effectiveness of falls prevention interventions for older community-dwelling Australians. Australian and New Zealand Journal of Public Health. 2012; 36(3):241-248

11. Craig J, Murray A, Mitchell S, Clark S, Saunders L, Burleigh L. The high cost to health and social care of managing falls in older adults living in the community in Scotland. Scottish Medical Journal. 2013; 58(4):198-203

12. Curtis LA, Burns A. Unit costs of health and social care 2020. Canterbury. Personal Social Services Research Unit University of Kent, 2020. Available from: https://kar.kent.ac.uk/84818/

13. Davies J, Rae TC, Montagu L. Long-term benzodiazepine and Z-drugs use in England: a survey of general practice [corrected]. British Journal of General Practice. 2017; 67(662):e609-e613

14. de Gier N, Gorgels W, Lucassen P, Oude Voshaar R, Mulder J, Zitman F. Discontinuation of long-term benzodiazepine use: 10-year follow-up. Family Practice. 2010; 28(3):253-259

15. de Jong MR, Van der Elst M, Hartholt KA. Drug-related falls in older patients: implicated drugs, consequences, and possible prevention strategies. Therapeutic Advances in Drug Safety. 2013; 4(4):147-154

16. de Joode SGCJ, Kalmet PHS, Fiddelers AAA, Poeze M, Blokhuis TJ. Long-term functional outcome after a low-energy hip fracture in elderly patients. Journal of Orthopaedics and Traumatology. 2019; 20(1):20

17. Department of Transport. Reported road casualties in Great Britain: notes, definitions, symbols and conventions - 2017. 2017. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/743853/reported-road-casualties-gb-notes-definitions.pdf Last accessed: 06 August 2021.

18. Department of Transport. Road safety data. 2021. Available from: https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data Last accessed: 06 August 2021.

19. Department of Transport. Statistical data set - Reported road accidents, vehicles and casualties tables for Great Britain. 2021. Available from: https://www.gov.uk/government/statistical-data-sets/reported-road-accidents-vehicles-and-casualties-tables-for-great-britain Last accessed: 06 August 2021.

20. Donnelly K, Bracchi R, Hewitt J, Routledge PA, Carter B. Benzodiazepines, Z-drugs and the risk of hip fracture: A systematic review and meta-analysis. PloS One. 2017; 12(4):e0174730-e0174730

21. Gray A, Clarke P, Wolstenholme J, Wordsworth S. Applied methods of cost-effectiveness analysis in healthcare. Oxford University Press,. 2011.

22. Griffin XL, Parsons N, Achten J, Fernandez M, Costa ML. Recovery of health-related quality of life in a United Kingdom hip fracture population. The Warwick Hip Trauma Evaluation--a prospective cohort study. Bone Joint J. 2015; 97-b(3):372-382

23. Haentjens P, Magaziner J, Colón-Emeric CS, Vanderschueren D, Milisen K, Velkeniers B et al. Meta-analysis: excess mortality after hip fracture among older women and men. Annals of Internal Medicine. 2010; 152(6):380-390

24. Health & Social Care Information Centre. Health survey for England - 2005, health of older people. 2007. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2005-health-of-older-people Last accessed: 06 August 2021.

25. Hopkins RB, Pullenayegum E, Goeree R, Adachi JD, Papaioannou A, Leslie WD et al. Estimation of the lifetime risk of hip fracture for women and men in Canada. Osteoporosis International. 2012; 23(3):921-927

26. Kragh A, Elmståhl S, Atroshi I. Older adults' medication use 6 months before and after hip fracture: a population-based cohort study. Journal of the American Geriatrics Society. 2011; 59(5):863-868

27. Landeiro F, Mughal S, Walsh K, Nye E, Morton J, Williams H et al. Health-related quality of life in people with predementia Alzheimer's disease, mild cognitive impairment or dementia measured with preference-based instruments: a systematic literature review. Alzheimer's Research & Therapy. 2020; 12(1):154

28. Leal J, Gray AM, Prieto-Alhambra D, Arden NK, Cooper C, Javaid MK et al. Impact of hip fracture on hospital care costs: a population-based study. Osteoporosis International. 2016; 27(2):549-558

29. Lucchetta RC, da Mata BPM, Mastroianni PC. Association between Development of Dementia and Use of Benzodiazepines: A Systematic Review and Meta-Analysis. Pharmacotherapy. 2018; 38(10):1010-1020

30. Matthews FE, Stephan BCM, Robinson L, Jagger C, Barnes LE, Arthur A et al. A two decade dementia incidence comparison from the Cognitive Function and Ageing Studies I and II. Nature Communications. 2016; 7(1):11398

31. Miller NS, Mahler JC. Addiction to and dependence on benzodiazepines: Diagnostic confusion in clinical practice and research studies. Journal of Substance Abuse Treatment. 1991; 8(1-2):61-67

32. Moriarty F, Cahir C, Bennett K, Fahey T. Economic impact of potentially inappropriate prescribing and related adverse events in older people: a cost-utility analysis using Markov models. BMJ Open. 2019; 9(1):e021832

33. Morin CM, Bastien C, Guay B, Radouco-Thomas M, Leblanc J, Vallieres A. Randomized clinical trial of supervised tapering and cognitive behavior therapy to facilitate benzodiazepine discontinuation in older adults with chronic insomnia. American Journal of Psychiatry. 2004; 161(2):332-342

34. Morin CM, Belanger L, Bastien C, Vallieres A. Long-term outcome after discontinuation of benzodiazepines for insomnia: a survival analysis of relapse. Behaviour Research and Therapy. 2005; 43(1):1-14

35. National Institute for Clinical Excellence. Dementia: assessment, management and support for people living with dementia and their carers. NICE guideline NG97. London. National Institute for Health and Care Excellence, 2018. Available from: https://www.nice.org.uk/guidance/ng97/chapter/Recommendations#managing-non-cognitive-symptoms

36. National Institute for Clinical Excellence. National Institute for Clinical Excellence: Our charter. 2021. Available from: https://www.nice.org.uk/about/who-we-are/our-charter Last accessed: 30/06/2021.

37. National institute for Clinical Excellence. Our principles: the principles that guide the development of NICE guidance and standards. 2021. Available from: https://www.nice.org.uk/about/who-we-are/our-principles Last accessed: 30/06/2021.

38. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [updated 2020]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

39. National Institute for Health and Care Excellence. Falls in older people: assessing risk and prevention. Clinical guideline CG161. London. National Institute for Health and Care Excellence, 2013. Available from: https://www.nice.org.uk/guidance/cg161

40. National Institute for Health and Care Excellence. Falls in older people: assessing risk and prevention. Clinical guideline CG161. Appendix K: Full health economic report. London. National Institute for Health and Care Excellence, 2013. Available from: https://www.nice.org.uk/guidance/cg161/evidence/cg161-appendix-k-full-health-economic-report-2013-pdf-6841214714

41. Neutel CI, Patten SB. Risk of suicide attempts after benzodiazepine and/or antidepressant use. Annals of Epidemiology. 1997; 7(8):568-574

42. NHS Business Services Authority. Prescription cost analysis (PCA) data - Changes to prescription cost analysis (PCA) monthly administrative data. 2021. Available from:

https://www.nhsbsa.nhs.uk/prescription-data/dispensing-data/prescription-cost-analysis-pca-data Last accessed: 06 August 2021.

43. NHS England and NHS Improvement. 2018/19 National Cost Collection Data Publication. 2021. Available from: https://www.england.nhs.uk/publication/2018-19-national-cost-collection-data-publication/ Last accessed: 30/06/2021.

44. Office for National Statistics. Dataset - Death registrations summary tables - England and Wales. 2018. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/deathregistrationssummarytablesenglandandwalesreferencetables Last accessed: 06 August 2021.

45. Office for National Statistics. Death registrations in England and Wales, summary tables: 2017. 2017. Available from: https://www.ons.gov.uk/releases/deathregistrationsinenglandandwalessummarytables2017 Last accessed: 06 August 2021.

46. Office for National Statistics. National Life tables: UK. 2020. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables Last accessed: 11 August 2021.

47. Oude Voshaar RC, Gorgels WJ, Mol AJ, van Balkom AJ, Mulder J, van de Lisdonk EH et al. Long-term outcome of two forms of randomised benzodiazepine discontinuation. British Journal of Psychiatry. 2006; 188:188-189

48. Oude Voshaar RC, Krabbe PF, Gorgels WJ, Adang EM, van Balkom AJ, van de Lisdonk EH et al. Tapering off benzodiazepines in long-term users: an economic evaluation. Pharmacoeconomics. 2006; 24(7):683-694

49. Prince M, Knapp M, Guerchet M, McCrone P, Prina M, Comas-Herrera A et al. Dementia UK - update. Alzheimer's Society, 2014. Available from: https://www.alzheimers.org.uk/sites/default/files/migrate/downloads/dementia_uk_update.pdf

50. Prudham D, Evans JG. Factors associated with falls in the elderly: a community study. Age and Ageing. 1981; 10(3):141-146

51. Rao A, Suliman A, Vuik S, Aylin P, Darzi A. Outcomes of dementia: Systematic review and meta-analysis of hospital administrative database studies. Archives of Gerontology and Geriatrics. 2016; 66:198-204

52. Richardson K, Bennett K, Kenny RA. Polypharmacy including falls risk-increasing medications and subsequent falls in community-dwelling middle-aged and older adults. Age and Ageing. 2015; 44(1):90-96

53. Taylor S, Annand F, Burkinshaw P, Greaves F, Kelleher M, Knight J et al. Dependence and withdrawal associated with some prescribed medicines: an evidence review. London. Public Health England, 2019. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/940255/PHE_PMR_report_Dec2020.pdf

54. Wittenberg R, Knapp M, Hu B, Comas-Herrera A, King D, Rehill A et al. The costs of dementia in England. International Journal of Geriatric Psychiatry. 2019; 34(7):1095-1103

Appendices

Appendix A: Forest plots

Figure 11: Cessation rate post intervention

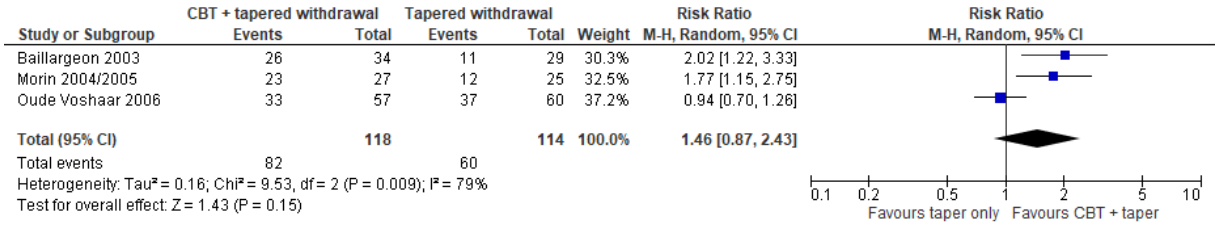
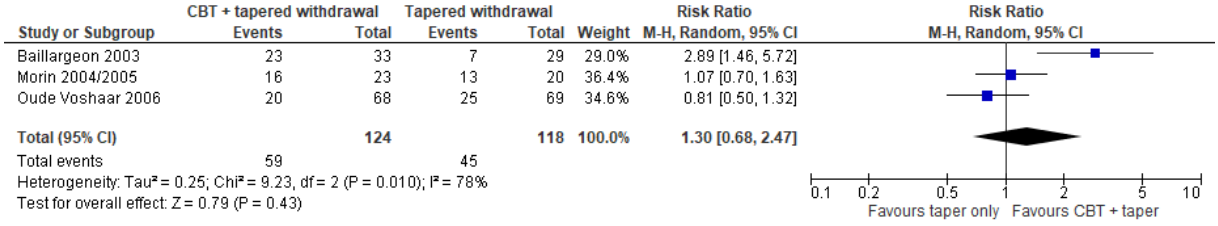


Figure 12: Cessation rate 12-15 months



Appendix B: SF-36 scores

Table 30: SF-36 scores

Intervention	Physical functioning	Social functioning	Role limitation (physical)	Role limitation (emotional)	Mental health	Vitality	pain	General health perception
Tapering off plus cognitive behavioural therapy								
Baseline	67 ± 26	61 ± 21	51 ± 40	63 ± 39	64 ± 18	55 ± 23	62 ± 27	58 ± 23
18 months	68 ± 26	68 ± 22	57 ± 44	67 ± 41	71 ± 17	63 ± 20	67 ± 26	62 ± 19
Tapering off alone								
Baseline	66 ± 25	64 ± 24	54 ± 42	69 ± 39	66 ± 20	57 ± 23	60 ± 24	52 ± 20
18 months	65 ± 26	64 ± 26	54 ± 42	76 ± 39	76 ± 39	61 ± 20	61 ± 27	57 ± 20
Usual care								
Baseline	68 ± 25	66 ± 19	69 ± 35	70 ± 42	64 ± 23	56 ± 24	66 ± 26	58 ± 22
18 months	72 ± 26	69 ± 19	76 ± 36	81 ± 29	81 ± 29	63 ± 24	69 ± 22	55 ± 22
Norm scores	**82**	**87**	**79**	**84**	**77**	**67**	**80**	**73**

Source: Oude Voshaar 2006[48]