National Institute for Health and Care Excellence

Osteoarthritis in over 16s: diagnosis and management

Cost-utility analysis: Oral, topical and transdermal pharmacological treatments

NICE guideline 226

Economic analysis report

October 2022

Final

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 All NICE guidance is subject to regular review and may be updated or withdrawn.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 Introduction

One of the key clinical issues explored in the osteoarthritis guideline update related to its pharmacological management since current practice varies widely nationally. It is estimated that most patients will receive paracetamol as a first line drug treatment. Remaining patients will be prescribed an opioid or a non-steroidal anti-inflammatory drug (NSAID). A growing proportion of patients are also being prescribed topical NSAIDs, though these can be purchased over the counter.

Based on the clinical evidence, the guideline committee decided not to recommend paracetamol for the management of osteoarthritis. It was agreed therefore that the opportunity cost of withdrawing paracetamol would be investigated via a health economic evaluation, looking specifically at the costs and Quality-adjusted life-years (QALYs) associated with each treatment alternative.

The previous guideline model included individual NSAID drugs with and without PPIs, combination formulations of NSAIDs with PPIs and paracetamol. The committee were specifically interested in drug classes and this model therefore considered the following:

- Paracetamol
- Oral NSAIDs plus PPIs
- Oral NSAIDs alone
- Topical NSAIDs
- Strong oral opioids
- Transdermal buprenorphine

Drug classes range from £20 - £100 per month at a minimum cost of the maximum dose. This guideline does not recommend long-term use, but the committee acknowledge there are instances where they are used long-term and would therefore accrue a lifetime cost.
The committee is concerned that there is a limited effectiveness for some of these drug classes. Furthermore, each drug class is thought to be associated with different harms. Therefore, there is potential for disinvestment in some drug classes for people with osteoarthritis. The committee noted that most of the oral drug classes will have a similar average cost to each other but noted that if one of the least expensive drug classes was not to be recommended and another slightly higher cost class of treatments were to be positively recommended, there could be a significant resource impact given the size of the osteoarthritis population. The committee noted that topical and transdermal medicines are significantly more expensive than oral medicines on the whole.
Given the different benefits and harms associated with each drug class as well as differences in drug costs, it is likely to be difficult to fully assess which is the most cost effective without an economic model quantifying the costs and QALYs associated with different adverse events.

2 Methods

2.1 Model overview

A cost-utility analysis was undertaken where lifetime quality-adjusted life years (QALYs) and costs from a 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.[70] An incremental analysis was undertaken.

2.1.1 Comparators

The following comparators were included in the analysis:
1. No treatment: patients do not have drug treatment for their osteoarthritis.
2. Paracetamol
3. Oral NSAIDs plus PPIs
4. Oral NSAIDs alone
5. Topical NSAIDs
6. Strong oral opioids
7. Transdermal buprenorphine:

Patients take one drug for a treatment duration of three months.

2.1.2 Population

The population of the analysis was adults with osteoarthritis.

In the base case analysis, the model was run using a starting age of 60 for men and women. This was based on the weighted average for age from the trials applicable to the base case analysis that were used to calculate age-related utilities. In the same manner, the ratio of the proportion of males to females were calculated as 0.35:0.65.

2.2 Approach to modelling

The aim of the model is to quantify the trade-off between improved quality of life due to intervention effectiveness and both cost and decreases in quality of life resulting from increased risks of adverse events. The costs and QALYs associated with each intervention were therefore calculated.

2.2.1 Model structure

A Markov model was constructed to calculate lifetime costs and QALYs for each comparator.

In a Markov model, a set of mutually exclusive health states are defined that describe what can happen to the population of interest over time. People in the model can only exist in one of these health states at a time. Possible transitions are defined between each of the health states and the probability of each transition occurring within a defined period of time (a cycle) is assigned to each possible transition.

Figure 1 illustrates the health states in the model and transitions between them in each cycle. A 3-monthly cycle length was used. A cohort of people entered the model for each individual intervention at which point they were exposed to the intervention-specific risk of adverse events.

The different adverse events can be broadly categorised into two: acute and chronic. The acute adverse events are transient and last for one cycle only. For those acute adverse events related to the gastro-intestinal system (bleeding, dyspepsia, symptomatic ulcer, constipation, nausea, vomiting) and central nervous system (vertigo) the person moves to the 'no treatment' state post-event.

The remaining acute adverse events (cardiovascular (CV), acute liver failure (ALF), acute kidney injury (AKI)) were set up as tunnel states where people move automatically to the post-event state in the following cycle (unless they die). This means that repeat cardiovascular events were not explicitly modelled, which was considered a reasonable simplification for modelling purposes.

Each adverse event was associated with a unique utility decrement, cost and probability of death. This model was run for 240 cycles (60 years) by which time everyone in the model was in the dead state, regardless of which treatment comparator they were allocated.

The model structure was the same for each comparator. Drug treatments were taken for one cycle only (3 months), and each treatment had its own treatment effectiveness (measured as change in EQ-5D), drug costs and probabilities of adverse events. This resulted in different total costs and total QALYs for each comparator. Comparing these results allowed us to identify whether each individual treatment was cost effective compared to no treatment and also how they compared against each other.

Figure 1. Model structure
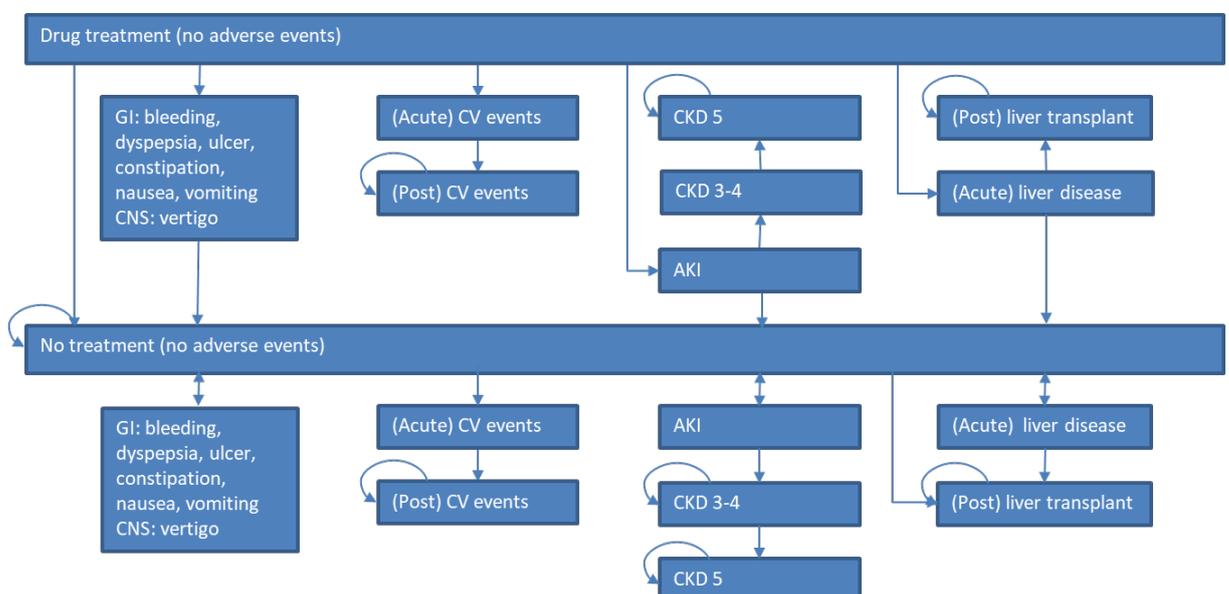
Abbreviations: AKI= acute kidney injury; CNS= central nervous system; CKD= chronic kidney disease; CV= cardiovascular; GI= gastrointestinal

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 for the base case and for each sensitivity analysis – and results were summarised.

To ensure the number of model runs was sufficient to account for random variation in sampling, we checked for stability in the incremental net health benefit gained for each intervention versus no treatment. The incremental net health benefit is a summary statistic based on incremental QALYs, incremental costs and the cost effectiveness threshold that highlights the change in population health resulting from the introduction of a new intervention. 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 10,000 runs.

Figure 2. Checking for convergence: incremental net health benefit versus no treatment at a threshold of £20,000 per QALY
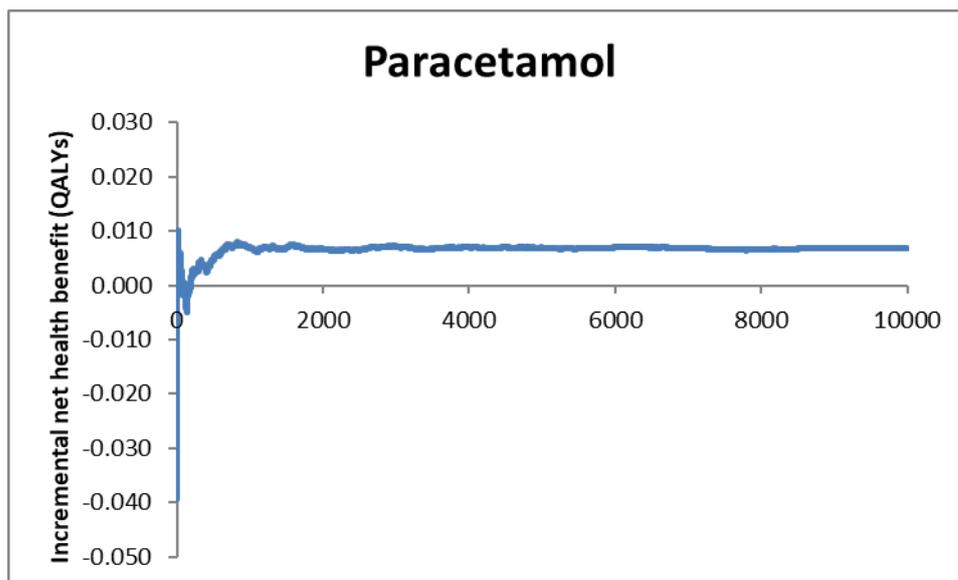
The way in which distributions are defined reflects the nature of the data, so for example event probabilities were given a beta distribution, which is bounded by 0 and 1, reflecting that the probability of an event occurring cannot be less than 0 or greater than 1. All of the variables that were probabilistic in the model and their distributional parameters are detailed in Table 1 and in the relevant input summary tables in section 2.3. 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
Probability of adverse events	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 having the event) Beta = (total number of patients) − (number of patients having the event)
Probability of first cardiovascular event	Dirichlet	Fitted to multinomial data. Represents a series of conditional distributions, bounded on 0–1 interval. Derived by the number of patients in the sample and the number of patients in a particular subgroup.
Odds ratios Hazard ratios Standardised mortality ratios Risk ratios	Lognormal	The natural log of the mean and standard error were calculated as follows: • Mean = $\ln(\text{mean cost}) - SE^2/2$ • SE = $[\ln(\text{upper 95\% CI}) - \ln(\text{lower 95\% CI})]/(1.96 \times 2)$ This formula includes a correction to ensure the mean generated in the probabilistic analysis will be the same as the reported mean. [7]
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 = $\text{mean}^2 \times [(1-\text{mean})/SE^2] - \text{mean}$ Beta = $\text{alpha} \times [(1-\text{mean})/\text{mean}]$
Utility decrements Costs Outcome measures (WOMAC, SF-36, VAS)	Gamma	Bounded at 0, positively skewed. Derived from mean and its standard error. Alpha and beta values were calculated as follows: • Alpha = $(\text{mean}/SE)^2$ • Beta = SE^2/Mean

Abbreviations: 95% CI= 95% confidence interval; SE= standard error; SF-36= 36-item short-form survey; VAS= visual analogue scale; WOMAC= Western Ontario and McMaster Universities Osteoarthritis Index.

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

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

• The cost-effectiveness threshold

• Drug prices

• Unit costs and resource use for healthcare (except for cardiovascular and hip fracture costs where the source literature reported standard errors)

• Excess mortality resulting from acute non-bleeding gastrointestinal events and vertigo since these were zero in the base case.

2.3   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 description of sources,

calculations and rationale for selection can be found below. A full list of inputs and their probabilistic parameters is provided in Appendix B.

2.3.1 Baseline adverse event probabilities

2.3.1.1 Gastrointestinal bleeds

The probability of gastrointestinal (GI) bleeds in the no treatment arm was taken from all relevant trials between oral NSAIDs alone and no treatment, oral NSAIDs with PPIs and no treatment and oral NSAIDs with PPI and oral NSAIDs alone. [11, 14, 15, 21, 37-41, 48, 49, 53, 54, 58, 80-84, 88, 89, 92, 100]

2.3.1.2 Gastrointestinal non-bleeding events

The non-bleeding gastrointestinal symptoms most frequently observed in clinical trials with strong oral opioids and transdermal buprenorphine were constipation, nausea and vomiting. These were therefore grouped together as associated with opioids for the model. The baseline probabilities for each were calculated by taking a weighted average of the number of patients reporting the events in the no control arm over a 3-month period (see Table 2). Only those opioid-related trials that were listed in the clinical review for the relevant gastrointestinal non-bleeding events were included.

Table 2. Number of patients in the control arm of opioid-related clinical trials reporting adverse events of constipation, nausea, vomiting or vertigo

Study	Duration	N (control)	Constipation	Nausea	Vomiting	Vertigo
Afilalo 2010	3 months	337	22	23	11	16
Serrie 2017	3 months	337	31	21	13	29
Breivik 2010	6 months	91	5	10	2	18
Total	-	765	58	54	26	63
Percentage	**-**		**7.6%**	**7.1%**	**3.4%**	**8.2%**

There were two adverse events commonly reported with oral NSAIDs, dyspepsia and symptomatic ulcer. Baseline probabilities were calculated in the same manner as described above with opioids but using the control arm of trials comparing NSAIDs with placebo (Table 3).

Table 3. Number of patients in the control arm of oral NSAID-related clinical trials reporting adverse events of dyspepsia or symptomatic ulcer

Study	Duration	N (control)	Dyspepsia	Symptomatic ulcer
Bocanegra 1998	3 months	91	35	3
Amundsen 1983	3 months	52	1	
Andelman 1983	3 months	10	0	
Anonymous 1983	3 months	599	0	
Baerwald 2010	3 months	30	0	
Bakshi 1991	3 months	106	0	
Bensen 1999	3 months	203	1	0
Couto 2018	3 months	409	3	0
Dore 1995	3 months	86	0	
Essex 2012	3 months	66	1	
Essex 2014	3 months	61	1	
Famaey 1976	3 months	20	6	
Fleischmann 2006	3 months	94	0	
Ghosh 2007	3 months	126	6	0
Golden 2004	3 months	155	0	0
Gordo 2017	3 months	79	2	
Hubault 1976	3 months	9	7	
Kageyama 1973	3 months	43	0	
Karakaya 1977	3 months	5	0	
Kivitz 2001b	3 months	218	0	
Kivitz 2004	3 months	208	0	
Leung 2002	3 months	56	16	0
Lopez sanchez 1983	3 months	10	0	
Lund 1998	3 months	137	4	
Makarowski 2002	3 months	117	0	0
Paul 2009	3 months	141	0	
Pincus 2004	3 months	289	5	0
Puopolo 2007	3 months	111	0	
Sandelin 1997	3 months	82	0	
Schmitt 1999	3 months	56	0	0
Schnitzer 2011A	3 months	257	0	0
Schnitzer 2011 B	3 months	416	11	0
Schubiger 1980	3 months	34	0	8
Scott 2000	3 months	303	0	0
Sheldon 2005	3 months	382	0	0
Tannenbaum 2004	3 months	243	9	0
Wasselman 1984	3 months	14	0	

Study	Duration	N (control)	Dyspepsia	Symptomatic ulcer
Wiesenhutter 2005	3 months	104	0	
Yocum 2000	3 months	157	0	
Total	-	5,579	108	11
Percentage	**-**	**-**	**1.9%**	**0.35%**

Note that the probability of symptomatic was based on only those trials that reported for ulcer (number of participants= 3,137)

2.3.1.3   Cardiovascular events

There were eight health states linked to cardiovascular (CV) events in the model: stable angina (SA), unstable angina (UA), myocardial infarction (MI), transient ischaemic attack (TIA), non-fatal stroke, death resulting from CHD and death resulting from CVD. These were the health states applied to CV events in the model for the NICE hypertension guideline and the committee agreed that they would also be relevant for a model in the osteoarthritis population.

The three-monthly transition probabilities for each individual CV event were calculated following the methodology used in the hypertension guideline model [65], taking into account:

* The overall rate of CV events [97]

* The relative distribution of the individual CV events that make up this overall rate, which were taken from the QRISK2

* Information about CV events not included in QRISK2 (i.e., heart failure)[23]

Transition probabilities therefore changed over time because of two reasons: firstly, within the overall risk of any CV event, the breakdown of which event is more or less likely varied with age and sex. Secondly, the overall risk of any CV event increased with each cycle as the cohort aged, as age is the main determinant of risk (Table 4).

Table 4. Incidence rates of cardiovascular events

Age	Total incidence rate of CV events	HF incidence rate
Male		
45-54	0.0042	0.0003
55-64	0.0137	0.0017
65-74	0.0243	0.0039
75-84	0.0375	0.0098
85+	0.0426	0.0168
Female		
45-54	0.0016	0.0001
55-64	0.0066	0.0007
65-74	0.0124	0.0023
75-84	0.0234	0.0059
85+	0.0329	0.0096

Annual incidence rates were converted to 3-monthly probabilities using the following formula:

	Where
$Transition\ probability\ (P) = 1 - e^{-rt}$	r = selected rate t = cycle length (0.25 years)

Calculating the cardiovascular event-specific risk element

The relative distribution of first CV events that are included in the QRISK2 tool (SA, UA, MI, TIA, stroke, CV death), were based on the same source as the Hypertension guideline models: Ward 2007.[97] A rate for each event was calculated using this distribution combined with an overall age/sex specific cardiovascular event rate (Ward 2007[96]).

Heart failure is not included in the QRISK2 tool but was included in the model. The incidence rate of heart failure was also taken from the Lipids model (Cowie 1999[23]).

The final probabilities for CV events are presented in Table 5.

Table 5: 3-month probability of cardiovascular events including heart failure

Age (years)	SA	UA	MI	TIA	Stroke	HF	CVD death
Male							
45-54	0.00032	0.00011	0.00031	0.00006	0.00014	0.00007	0.00003
55-64	0.00112	0.00024	0.00059	0.00030	0.00071	0.00029	0.00016
65-74	0.00130	0.00050	0.00105	0.00061	0.00164	0.00059	0.00038
75-84	0.00179	0.00076	0.00151	0.00075	0.00321	0.00059	0.00075
85+	0.00228	0.00102	0.00198	0.00017	0.00373	0.00059	0.00087
Female							
45-54	0.00013	0.00005	0.00003	0.00006	0.00009	0.00001	0.00002
55-64	0.00054	0.00019	0.00013	0.00026	0.00038	0.00006	0.00009
65-74	0.00107	0.00023	0.00029	0.00029	0.00089	0.00012	0.00021
75-84	0.00118	0.00030	0.00071	0.00043	0.00223	0.00047	0.00053
85+	0.00123	0.00028	0.00084	0.00081	0.00381	0.00035	0.00090

Abbreviations: HF= heart failure; MI= myocardial infarction; SA= stable angina; TIA= transient ischaemic attack; UA= unstable angina

2.3.1.4 Renal adverse events

Acute kidney injury

The incidence of acute kidney injury was taken from a study by Ali 2007.[2] This was a retrospective cohort study in Scotland based on 523,390 people from the area of Grampian. The reported incidence rate was 2,147 cases per million person years, which was subsequently converted to a 3-month probability.

Probability of progressing from AKI to stage 3-4 CKD

The clinical evidence is limited regarding the progression from AKI to stage 3-4 CKD. Arshad 2020 reported that 19% of a cohort of patients with AKI went on to develop CKD in one year.[4] This was a one-year retrospective follow-up study of patients with septic acute kidney injury that determined in what proportion of patients the injury resolved and what proportion it progressed to chronic kidney injury This figure was converted to a 3-month probability and used in the model.

Probability of progressing from CKD stage 3-4 to CKD stage 5

The baseline transition probability associated with the progression of CKD stage 3-4 to stage 5 CKD was taken from a ten-year cumulative incidence rate reported in a cohort study of 3,047 patients by Eriksen 2006. The patients in this study had an average eGFR of 55.1. While this is higher than expected in patients with stage 4 CKD, it was utilised in the model in the absence of data in patients with stage 4 CKD. The study reported the rate of progression

in three different age groups: <69, 70-79, >80 years. It was therefore possible to define the three-month probability of progressing from stage 3-4 CKD to stage 5 CKD as an age-group dependant variable (Table 6) using the formula:

$$\frac{-LN(1 - 10\ year\ rate)}{40}$$

This gave the 3-month rate which was then converted to a probability by exponentiating the rate:

$$1 - EXP(rate) * 1$$

In order to incorporate this age dependant variable, each probability is repeated three times for each age category.

Table 6: Age dependant disease specific progression from Stage 3-4 to stage 5 CKD

Age category (years)	10-year cumulative incidence	3-month probability
<69	0.07	0.0018
70-79	0.04	0.0010
>79	0.03	0.0008

Abbreviations:

2.3.1.5 Hepatic adverse events

Acute liver failure

The incidence of acute liver failure was taken from a study by Weiler 2020[99], which was based on statutory health insurance data from the largest health insurance provider in Germany between January 2014 and December 2018. The incidence of acute liver failure was presented as a distribution by gender and across decades in age, for which data from 60 years onwards was extracted for use in the model. The incidence per 100,000 years as reported in the study is presented below in Table 7 (note: these figures were manually read and interpreted from a graph).

Table 7. Incidence of ALF per 100,000 years, by age and gender

Age category (years)	Male	Female
60-69	0.75	0.73
70-79	1.2	1.15
80-89	1.25	1.57
90-99	0.75	1.91
100-109	0.00	1.99

These rates were converted to a 3-month probability for the model. It was assumed that the probabilities after age 110 years remained constant.

The acute liver failure health state is an acute state meaning that patients remain for one cycle only. After that, there are three possibilities: recovery and transition to the 'no treatment' state, liver transplant and subsequent transition to the post-liver transplant health or death.

Liver transplant

It is reported in a review of acute liver failure (Stravitz 2012) that approximately 25% of the people with acute liver failure will go on to have a liver transplant.[91] This probability was applied to the acute liver failure health state in the model, and after factoring in mortality resulting from the procedure, the remaining patients transfer to the post-liver transplant health state, where they remain for the duration of the model time horizon.

2.3.1.6 Central nervous system adverse events

Vertigo was a commonly reported adverse event of the central nervous system with strong oral opioids and transdermal buprenorphine. The baseline probability was calculated by taking a weighted average of the number of patients reporting the events in the control arm over a 3-month period (see Table 2).

2.3.2 Treatment effects - adverse events

The model focused on adverse events associated with treatment along with treatment efficacy (see section 2.3.4). Adverse events data associated with drug treatment were taken from the clinical review and applied to each cycle of the model for the relevant drug treatment (see Table 8 below).

Table 8. Risk ratio of adverse events versus no treatment

	Paracetamol	Oral NSAIDs plus PPI	Oral NSAIDs alone	Topical NSAIDs	Strong opioids	Transdermal buprenorphine
Gastrointestinal (bleeding or perforation)	1.00	2.32 (0.81 to 9.21)	2.99 (2.13 to 4.19)	1.00	1.00	1.00
Gastrointestinal (non-bleeding or perforation)	1.00	1.00	1.17 (0.39 to 3.57)	1.00	1.63 (0.80 to 3.28)	2.26 (1.54 to 3.30)
Cardiovascular	1.00	2.51 (0.73 to 8.59)	1.15 (0.84 to 1.56)	1.00	1.00	1.00
Hepatic	5.97 (2.30 to 15.50)	1.00	1.00	1.00	1.00	1.00
Renal	1.00	2.67 (0.50 to 20.50)	1.96 (0.18 to 20.80)	1.00	1.00	1.00
Central nervous system	1.00	1.00	1.00	1.00	1.93 (1.67 to 2.24)	2.48 (1.55 to 3.96)

2.3.3 Mortality

2.3.3.1 General population mortality

In each 3-month cycle, people in the no treatment arm in the model are at risk of death from non-cardiovascular causes and so a proportion will transition to the dead state.

Transition probabilities for the no treatment arm were based on the Office of National Statistics (ONS) life tables for England 2018-20.[75] The proportion of deaths that are non-circulatory were also taken from the ONS and applied to the mortality rates to determine the non-CV mortality rate by age and gender.

2.3.3.2 Adverse event mortality

Table 9. Excess mortality for acute adverse events

Event	Data	Source	Probability distribution
GI bleeds	9.34%	Roberts 2012[77]	Beta
Constipation	0%	Assumed	Fixed[a]
Nausea	0%		Fixed[a]
Vomiting	0%		Fixed[a]
Vertigo	0%		Fixed[a]
Angina stable)	Age- and sex-specific.	The distribution of CV mortality was attributed to individual health states using data from QRISK2. This was subsequently applied to the incidence of age- and sex-specific CV events, which were taken from Ward 2005. Rates were then converted to 3-month probabilities.	Gamma
Angina (unstable)			
Myocardial infarction			
Stroke			
Transient ischaemic attack			
Heart failure			
Acute kidney injury	18.2%	James 2010[46]	Beta
Acute liver failure (ALF)	52%	Bernal 2013[12]	Beta
Liver transplant resulting from ALF	23.85%	Dawwas 2007[30]	Beta
Vertigo	0%	Assumed to have no effect of general population mortality.	Fixed[a]

Abbreviations: GI= gastrointestinal
(a) Probabilities were fixed as these were zero values

Table 10. Mortality relative risks associated with chronic health states

Input	Data	Source	Probability distribution
Angina (stable)	1.95	Age-adjusted relative risk for death from any cause in men with angina (compared to men free from clinical CHD). 16-year follow-up. Swedish general population sample. Rosengren 1998[78]	Log normal
Angina (unstable)	2.19	Weighted average of SMRs for UA/NSTEMI 1 year in those alive at 6 months with and without new MI. UA/NSTEMI NICE guideline.[59]	Log normal
Myocardial infarction	2.68	Average of SMRs for men and women. All-cause mortality after first non-fatal MI compared to that expected in the general population. Danish population. Up to 15 years follow up. Bronnum-Hansen 2001[18]	Log normal
Stroke	1.4	Average of SMRs for men and women. All-cause mortality after first on-fatal stroke compared to that expected in the general population. Danish population. Up to 15 years follow up. Bronnum-Hansen 2001[17]	Log normal
Transient ischaemic attack	2.72	Risk ratio for mortality in people with TIA compared to that expected in those without TIA (age and sex matched). UK population. Mean of 3.7-year follow-up. Oxfordshire Community Stroke Project. Dennis 1990[32]	Log normal
Heart failure	2.2	Based on the SMR used for the preserved ejection fraction heart failure (HF-PEF) population, where annual mortality from a trial with an average of 4 year follow up was compared to the general population annual mortality for the same age group to derive a crude SMR.[67]	Log normal
Chronic kidney disease (CKD) stage 3-4	Age-dependent SMRs	Eriksen 2006[34]	Log normal
Severe chronic kidney injury (CKD) stage 5	Age-dependent SMRs	Villar 2007[93]	Log normal
Liver transplant	2.3	Dawwas 2007[30]	Log normal

Mortality associated with gastrointestinal bleeding

The mortality associated with gastrointestinal bleeding was taken from the previous osteoarthritis model and is based on a study by Roberts 2012.[77] This was a study that reported on prognosis following upper gastrointestinal bleeds over 3 months in people of all ages.

Mortality associated with non-bleeding gastrointestinal events and vertigo

It was assumed there was no excess mortality associated with the following adverse events: dyspepsia, symptomatic ulcer, constipation, nausea, vomiting and vertigo. The model only factored in a disutility and a cost of treatment for these events.

Mortality associated with cardiovascular events

Patients in any of the CV adverse event states are at a higher risk of mortality than then general population. This risk was implemented by applying relevant standardised mortality ratios (SMRs) to the 3-monthly age-dependant general population mortality probabilities (all-cause mortality). SMRs were identified from the NICE hypertension in adults guideline and can be seen in Table 10. The SMRs were applied to both the acute event states and post event states.

Mortality associated with renal events

The mortality from acute kidney injury is taken from the study by James 2010.[46] This was a large retrospective cohort study of 14,782 adults undergoing coronary angiography, of which, 1,420 patients had contrast induced AKI. The study divided AKI into three stages and reported the probability of death for each stage. A weighed average of the probability of death in all three stages (18.2%) was used in the model for AKI.

The mortality associated with CKD stage 3-4 is taken from the study by Eriksen 2006.[34] The study provides an age and sex-dependent standardised mortality ratio (SMR) that are presented in Table 11. SMRs were then multiplied by the age dependant mortality from the life tables (standard UK mortality rates by age) provided by the Office of National Statistics. Mortality is applied at each cycle prior to the transition probabilities for the other health states

Table 11. Age dependent SMRs in stage 3-4 CKD for men and women by age

Age category (years)	Men (SMR)	Women (SMR)
<69	3.6	2.7
70-79	2.4	1.8
>79	2.3	2.1

Abbreviations: SMR= standardised mortality ratio

Death from stage 5 CKD was taken from Villar 2007[93], where age-dependant SMRs were reported, which are presented in Table 12.

Table 12: Age dependant standardised mortality ratios for Stage 5 CKD

Age Category (years)	Men (SMR)	Women (SMR)
18-64	8.88	13.86
>65	4.88	7.96

Abbreviations: SMR= standardised mortality ratio

Mortality associated with hepatic events

The mortality associated with acute liver disease was reported as 52% in Stravitz 2012[91], which was applied in the model.

The mortality for the acute and post-liver transplant states for acute liver failure were taken from Dawwas 2007.[30] This was a multicentre cohort study based on two national databases of all adults who underwent a first and single liver transplant in the UK, Ireland or the USA between March 1994 and March 2005. Five-year mortality after liver transplant for acute liver failure was reported as 34.1% in the UK and Ireland. Annual figures for mortality post-transplant were also presented in a graph over that five-year period. For the purposes of the model, it was assumed that mortality in year 1 was equivalent to the acute liver transplant

health state and mortality between years 2-5 was equivalent to the post-liver transplant health state.

The probability of death in the first year was estimated as 26% based on a reading of the graph in Dawwas 2007. The probability of death in years 2-5 was 10.9%. This probability was transformed into a rate and compared to the average mortality rate in the general population with the corresponding age. This enabled the calculation of the standardised mortality ratio (SMR) that was used in the model to calculate the probability of death in the chronic state. The mortality in the first year from the graph was transformed into a 3-month mortality assuming that all the acute deaths occur in the first three months whereas in the other months, the mortality was as for the chronic state.

2.3.4 Utilities

Quality Adjusted Life Years (QALY) are the key outcome of the model. However, there are few trials reporting utility scores, which would be needed to calculate QALYs. There are two key areas for which quality of life data is important in the model. These are:

- efficacy of the different treatments (2.3.4.1 to 2.3.4.4) and
- the adverse events profile of the different treatments (0).

2.3.4.1 Calculating treatment efficacy

In the model base case, EQ-5D scores for all interventions were mapped from SF-36 or WOMAC scores. For studies that reported SF-36 data, the mean for each summary/domain score along with the standard deviation were extracted for the baseline and any subsequent time points, for both the intervention and control groups.

For studies that reported WOMAC scores, both the total WOMAC score and the subscale scores along with the standard deviations were extracted for the baseline and any subsequent time points, for both the intervention and control groups. One study did not report the WOMAC stiffness subscale score, so it was instead calculated by subtracting the other subscale scores from the total WOMAC score. For that study, the standard error was assumed to be 20% of the subscale score.

VAS pain scores were also extracted from studies that were identified during the clinical review and these were incorporated into utility calculations for a sensitivity analysis. If a study did not report the baseline score or any other relevant data for the mapping algorithm such as mean age/sex profile of participants, a weighted average of other studies where this was reported was used instead.

2.3.4.2 Mapping algorithms

Mapping SF-36 summary scores to EQ-5D

An algorithm by Price 2019 was chosen in the base case to convert from SF-36 physical and mental component scores (PCS and MCS) to EQ-5D.[76] It focuses on an osteoarthritis population and uses a UK population tariff for EQ-5D scores, thereby meeting the NICE reference case. It has other advantages to such as many observations underpinning the results (N=19,410 observations from 2,201 individuals) as well as the availability of a co-variance matrix, which is useful when making its algorithm probabilistic. The advantage of a co-variance matrix is that it enables interrelation between individual variables.

A scenario analysis used an algorithm by Lawrence 2004 instead.[50] Although this algorithm uses a general US population tariff to derive EQ-5D values, it also has many observations underpinning the results (14,580 individuals, of which 7,313 were selected randomly for the analytic sample and 7,267 were reserved for validation of the mapped scores) as well as a co-variance matrix. Three regression models were used during mapping. Although the 6-variable model had the best reported goodness of fit (R^2=0.628) and predictive ability (Mallow's C_p= 22.1) of the three, the 2- and 3-variable models performed better in predicting the EQ-5D scores across a range of disease areas. Since predicted scores were virtually identical between the 2- and 3- variable models, the authors used the 2-variable model for subsequent analysis, and this was the only model for which the variance-covariance matrix was published. For these reasons the 2-variable model was used here during mapping.

A second scenario analysis was chosen in which an algorithm by Maund 2012 was selected.[57] This was a systematic review and cost-utility analysis, that derived the utilities by creating a regression to map from SF-36 summary scores to the EQ-5D. The dataset used to generate the regression was the SAPPHIRE trial (2008) [98], which was in a population with rotator cuff disease (N = 200). The algorithm was based on a regression model using individual-level data at 1,3 and 12 months. This dataset was preferred to the 3-month dataset as the explanatory power and fit was better. There were five models to choose from, of which, 3 were ordinary least squares (OLS) models, one was a tobit model and one was a CLAD model. Of the OLS models, model 3 had the highest explanatory fit (adjusted R^2=0.4284) as well as the closest predicted EQ-5D score to the actual EQ-5D score and was therefore chosen for mapping.

It should be noted that although the Price and Lawrence algorithms were originally intended to convert SF-12 to EQ-5D, it is possible to map from either the SF-12 or SF-36 summary scores to EQ-5D-3L since both utilise the same summary scores on the same scale.

Mapping SF-36 domain scores to EQ-5D

SF-36 domain scores were mapped to EQ-5D-3L (UK tariff) utility scores using a mapping algorithm from Ara and Brazier 2008.[3] Model 5 was chosen as it was one of two models that accurately predicted the actual EQ-5D score, and of the two, it had the higher R^2. This algorithm only requires five of the eight domains: physical functioning, social functioning, role limitation (emotionally), mental health, general health as well as the average age of the cohort.

Mapping WOMAC to EQ-5D

WOMAC scores were mapped to the EQ-5D-3L (UK tariff) in the base case using the regression model from Wailoo 2014[95], while a sensitivity analysis explored the effect of using regression model E from Barton 2008.[9] The Wailoo 2014 model was based on 7,072 observations from 1,768 patients recruited in a registry study from 15 hospitals across Spain who were either scheduled to undergo primary joint replacement surgery due to knee/hip osteoarthritis or had received postoperative management. Of the available models, the five-class mixture model was preferred due to its superior summary measures of fit (MAE, RMSE, AIC and BIC). This model used the distribution of the EQ-5D UK value set to predict EQ-5D as a function of the WOMAC subscale scores.

The study by Barton 2008 mapped total WOMAC scores to EQ-5D using responses from individuals taking part in the Lifestyle Interventions for Knee Pain (LIKP) study. Inclusion criteria for the LIKP study were knee pain on most days over the past month, age greater than or equal to 45 and a BMI greater than 28 kg/m^2. The EQ-5D and WOMAC scores were completed at baseline by 348 individuals and 259 individuals further completed responses at 6,12 and 24 months. Five models were developed, of which model E had the highest adjusted R^2 (0.313) and the lowest MAE and RMSE (0.129 and 0.180, respectively) and was therefore selected for mapping purposes.

A second sensitivity analysis used an algorithm by Price 2019 to assess the impact of the above-mentioned algorithms on the cost effectiveness results.[76] This algorithm is based on a trial of patients with chronic pain of the knee suggestive of osteoarthritis in the UK (N=261) and uses a UK tariff for conversions to EQ-5D. However, it should be noted that this algorithm has not been externally validated, unlike the Wailoo and Barton algorithms. [47]

Mapping pain visual analogue scales (VAS) to EQ-5D

For studies that did not report WOMAC, SF36 or SF12 but did report VAS pain, change scores and baseline scores were extracted along with standard errors. Some studies did not report a baseline score, and here, it was assumed to be 6.5 since this was the average score where studies did report a baseline.

The pain scores and their confidence intervals were mapped onto the EQ-5D-3L (UK tariff) using the regression by Maund 2012. Note that the regression used by Maund was based on a dataset using the VAS on a 0-100 scale. Since the majority of trials reported VAS pain using the 0-10 scale, all scores were extracted in accordance with a 0-10 scale and then multiplied by 10 to convert them to the 0-100 scale.

Maund 2012 was a systematic review and cost-effectiveness analysis that created a regression to map from VAS pain to the EQ-5D. The dataset used to generate the regression was the SAPPHIRE trial (2008). The analysis with the largest population was used, which was the analysis using patient-level data reported at 1, 3 and 12 months (n= 491, of which, 295 were in the estimation data set (60%), and 196 in the validation data set). The OLS model that included the squared VAS interaction term was chosen as equations that include interaction terms are generally considered more reliable. Although other models were also available like TOBIT models, they did not report the R squared statistic which was needed (see section 2.2.2 for further details).

The model goodness of fit was poor, with an R squared of 0.1. However, this was the only mapping study identified that mapped VAS pain to EQ-5D without the need for further data. This approach to calculating EQ-5D score for a cost effectiveness model has been used elsewhere, most notably in a large acupuncture study.[55] In this analysis, a variance adjustment method was utilised to account for the additional uncertainty expected to arise during mapping, represented by the low R squared. This method is explained in more detail below.

Accounting for uncertainty in the regression weights

The coefficients in the mapping algorithms were themselves made probabilistic to account for uncertainties in the mapping equations. Various methods were used but they all drew values from a normal distribution.

Standard errors for the coefficients were reported with the Maund algorithms for SF-36 and VAS pain, so these were used to make the values probabilistic. They were not reported with the Barton algorithm, so they were assumed to be 20% of the point estimates. The Lawrence, Price SF-12 and Price WOMAC as well as the Wailoo algorithms all had variance/covariance matrices, and the Cholesky decomposition method was used to make the point estimates probabilistic, accounting for the covariance between the algorithm coefficients.

Accounting for uncertainty in mapping algorithms

It has been noted that the application of mapping algorithms, even probabilistically as described above, can underestimate uncertainty. The most obvious explanation for the variance underestimation of derived utilities is that there are important unmeasured predictors in most mapping algorithms. This leads to a relatively high degree of unexplained variance of utilities. In OLS based mapping algorithms, this is reflected as a relatively low R squared.

There were five OLS based mapping (Barton 2008 (WOMAC), Lawrence 2008 (SF-12/36), Maund 2012 (SF-12/36), Maund 2012 (VAS) and Ara & Brazier 2008 (SF-36) algorithms used in this model and a high level of unexplained variation was reported in all (that is, a relatively low R squared). To account for this, an additional variance component was included in the EQ-5D predictions.

Chan 2014 suggests methods that could be used to estimate the variance of mapped values, by accounting for a low R squared in OLS-based mapping algorithms.[20] Multiple methods are suggested, but some are only possible if patient-level data is available. One simple method that could be used to account for an artificially low variance of utilities because of a low R squared, is to inflate the variance of the derived utilities by a factor of 1/R squared. This estimator helps account for a low R squared but does not account for the uncertainty of the regression coefficients. This adjustment has also been used in other studies using a mapping algorithm for pain.[55]

This adjustment factor was applied to the variance of the mapped EQ-5D values for utilities mapped from WOMAC using the Barton algorithm (adjusted R squared = 0.313), from the SF-12/SF-36 using the Lawrence algorithm (R squared = 0.612) and Maund SF-12/36 algorithm (adjusted R squared = 0.428), form SF-36 using the Ara & Brazier algorithm (R squared = 0.5856) and from VAS using the Maund algorithm (R squared = 0.101).

2.3.4.3 Trials reporting treatment efficacy

Studies reporting treatment outcomes with paracetamol

The clinical review found one study that compared paracetamol to no treatment.[45] Treatment outcomes were reported using the WOMAC scale at baseline and 6 months, which were all extracted along with their associated standard deviations. The study did not report the WOMAC stiffness subscale score, but this was calculated by deducting the sum of the pain and physical function subscale scores from the WOMAC total score. The standard deviation for the WOMAC stiffness score was assumed to be 20% of the point estimate.

Studies reporting treatment outcomes with oral NSAIDs plus PPI

There were no studies identified during the clinical review that reported non-VAS pain outcomes. Therefore, the committee agreed it was reasonable to assume the change in utility would be identical to that reported for NSAIDs alone, since the addition of PPIs is intended to mitigate adverse events and is not to give a treatment benefit.

Studies reporting treatment outcomes with oral NSAIDs alone

Studies that reported visual analogue scale (VAS) pain scores were not included in the base case analysis. As a result, there were 4 studies identified that reported treatment outcomes with NSAIDs alone versus no treatment. Of these, two reported SF-36 summary scores[31, 81], one reported SF-36 domain scores[90], and the final study reported WOMAC scores.[89]

Baseline SF-36 summary scores were not reported in either of the two relevant trials, therefore these were assumed to be 30 for the physical component score on a scale of 0-100 and 50 for the mental component score on a scale of 0-100. These values were chosen as they aligned with what was reported in other trials in this model and in the guideline electroacupuncture

model (see separate report). Standard deviations were assumed to be 20% of the point estimates.

Studies reporting treatment outcomes with topical NSAIDs

There were eight studies in the clinical review that reported treatment outcomes for topical NSAIDs versus no treatment. Of the seven studies, two did not report WOMAC stiffness and WOMAC total scores[6, 8], which meant the reported outcomes could not be mapped to EQ-5D since all mapping algorithms require either the total WOMAC score or all three subscale scores. Of the six remaining studies [5 13, 16, 79, 89, 94], weighted averages of WOMAC scores were calculated based on the number of participants across the studies, which were then used in the calculations for conversion to EQ-5D.

One study did not report the mean age of patients[13], so it was assumed to be 61 in line with the model base case. Where total WOMAC scores were not reported, it was calculated by summing up the WOMAC pain, stiffness and physical function subscales scores. The standard deviation was assumed to be 20% of the point estimate.

Studies reporting on treatment outcomes with strong oral opioids

There were four studies reporting treatment outcomes over 3 months with strong oral opioids versus placebo that were used in the model. Of these, two reported changes in EQ-5D[1, 87], and the other two reported SF-36 summary scores.[31, 36]

The baseline EQ-5D score was not reported in either of the two relevant trials and was therefore assumed to be 0.5 since this was the mid-score.

Baseline SF-36 summary scores were also not reported in either of the two relevant trials, therefore these were assumed to be 30 for the physical component score on a scale of 0-100 and 50 for the mental component score on a scale of 0-100. These values were chosen as they roughly aligned with what was reported in other trials in this model and in the electroacupuncture model.

Standard errors were not reported and in all these cases were assumed to be 20% of the point estimates.

One study was identified that compared buprenorphine to no treatment. Treatment outcomes were reported using the WOMAC scale at baseline and 6 months.

It was assumed that the change scores at 3 months was identical to those reported at 6 months, therefore the mapped EQ-5D scores were directly applied to a single model cycle.

2.3.4.4   Base case treatment efficacy

The resulting EQ-5D scores to be used in the base case are presented in

Table **13** below.

24

Table 13: EQ-5D-3L mapped over time by randomised trial

Drug class	EQ-5D score	EQ-5D gain versus no treatment	Mapping algorithm used
No treatment	0.400[a]	-	-
Paracetamol	0.435	0.035	Wailoo 2014
Oral NSAIDs plus PPI	0.451[b]	0.051	-
Oral NSAIDs alone	0.451	0.051	Ara & Brazier 2008, Maund 2012, Wailoo 2014
Topical NSAIDs	0.502	0.102	Wailoo 2014
Strong oral opioids	0.409	0.009	Price 2019
Transdermal buprenorphine	0.461	0.061	Wailoo 2014

Abbreviations: NSAIDs= non-steroidal anti-inflammatory drugs; PPI= proton pump inhibitor
(a) Based on a weighted average of the baseline score in the no treatment arm across all trials
(b) Assumed identical to the value reported for NSAIDs alone in the absence of data

Age- and sex- specific utility values from the general population were used for the people in the model (Ara 2010).

Utility multipliers for people with osteoarthritis were calculated separately for each treatment and each gender as follows:

1. Mapping mean baseline SF-36 scores or WOMAC scores to EQ-5D using various mapping algorithms in the placebo arm for all relevant trials.
2. Taking from Ara 2010 the general population utility scores according to the average base case patient, which were 0.833 and 0.812 for males and females aged 61 years, respectively
3. Dividing the former by the latter

These multipliers (Table 14) were then applied to the general population utility scores to give age- and sex-specific utility values for people with each intervention. For example, for males in the no treatment arm, the following calculation was applied: 0.400/0.833= 0.480.

Table 14. Derivation of osteoarthritis utility multipliers

	Mean EQ-5D utility before age adjustment	Utility multiplier (males)	Utility multiplier (females)
No treatment	0.400	0.480	0.493
Paracetamol	0.435	0.523	0.536
Oral NSAIDs+PPI	0.451	0.550	0.565
Oral NSAIDs alone	0.451	0.550	0.565
Topical NSAIDs	0.502	0.603	0.619
Strong opioids	0.409	0.491	0.504
Transdermal buprenorphine	0.461	0.553	0.568

Abbreviations: NSAIDs= non-steroidal anti-inflammatory drugs; PPI= proton pump inhibitor

2.3.4.5 Adverse event utilities

Comparative utility scores for adverse events are important in the model because the adverse events associated with the different drugs are the key drivers of health effects as well as costs. However, data for the utility scores required were sparse, largely because of the time periods considered. Often utility scores are reported for adverse events without being specific about the time periods the utility scores relate to. This is important because a

utility score for a period after a MI will be very different one year after the event compared to one or two months after the event. For our model, we required average utility scores for the 3 months immediately after an event, as well as average scores for the longer term after an event.

In the absence of scores for nausea, vomiting and vertigo, the utility applied to the adverse event health state dyspepsia from the previous osteoarthritis guideline model was used as a proxy. The gastrointestinal bleeds health state was present in the previous model, so this utility score was carried over. Both utility scores were taken from a Canadian Coordinating Office for Health Technology Assessment (CCOHTA) survey reporting 3-month utility scores for various gastrointestinal adverse events associated with NSAIDs (Maetzel 2003).[56]

Utilities for CV events (and also for the post CV states) were taken from the NICE Hypertension model (2009).[67] Likewise, the utilities for renal-related health states were taken from the NICE contrast-induced acute kidney injury (CI-AKI) model (2013)[61], while utilities for liver transplant and its post-state were taken from the NICE non-alcoholic fatty liver disease (NAFLD) model (2016).[69] It was assumed that the utility of acute liver disease was the same as acute kidney disease in the absence of data.

The full list of utility multipliers for adverse events are shown in Table 15. These were multiplied by age- and sex-specific utility scores for the osteoarthritis population for the duration of the adverse event.

Table 15. Utility weights

Adverse event	Utility weight (1=OA, no complications)
Nausea	0.733
Vomiting	0.733
Gastrointestinal bleeds	0.459
Stable Angina (New Event)	0.808
Stable Angina (Post Event)	0.808
Unstable Angina (New Event)	0.770
Unstable Angina (Post Event)	0.808
MI (New Event)	0.760
MI (Post Event)	0.880
TIA (New Event)	0.900
TIA (Post Event)	0.900
Stroke (New Event)	0.628
Stroke (Post Event)	0.628
Heart Failure (New event)	0.683
Heart Failure (Post Event)	0.683
AKI	0.525
Moderate CKD	0.861
Severe CKD	0.798
ALF	0.525
Liver transplant	0.800
Liver transplant (Post event)	0.850
Vertigo	0.733

For the remaining adverse event health states, a utility decrement was applied instead to age-specific utility scores (Table 16).

Table 16. Utility decrements

Adverse event	Utility decrement
Constipation	-0.072

The utility decrement for constipation was taken from the NICE opioids in palliative care model (2012), which derived the value from a SF-36 data from a systematic review of constipation on quality of life in adults and children (Belsey 2010).[10] The SF-36 data were converted to EQ-5D using an algorithm published by Ara & Brazier 2008.[3]

The adverse events states for nausea, vomiting, vertigo and constipation were assumed to last 2 weeks after which people reverted to their age and sex-related utility score.

2.3.5   Resource use and costs

2.3.5.1   Drugs

The cost of the different drug treatments were obtained from the Drug Tariff[72], with doses based on Average Daily Quantities (ADQs) for the relevant indication taken from the online BNF.[64] A weighted average of drug class costs was used based on prescription usages data between 2020/21 released by the NHS.[73] The resultant drug costs per cycle are presented in

Table 17.

Table 17. Drug costs per model cycle

Drug treatment	Drugs included	Weighted average cost per day	Cost per cycle
No treatment	-	-	-
Drugs used to treat symptoms of osteoarthritis			
Paracetamol	Paracetamol alone	£0.26	£24.10
Oral NSAIDs	Aceclofenac, Celecoxib, Diclofenac, Diclofenac/misoprostol, Etodolac, Etorocoxib, Ibuprofen, Indometacin, Mefenamic acid Meloxicam, Naproxen, Naproxen/esomeprazole,	£0.19	£17.19
Topical NSAIDs	Diclofenac, Ibuprofen, Ketoprofen, Piroxicam	£0.31	£28.40
Strong oral opioids	Morphine, Oxycodone, Tapentadol, Tramadol,	£0.56	£51.23
Transdermal buprenorphine	Buprenorphine	£0.89	£81.64
Drugs used to prevent treatment side-effects			
Proton pump inhibitors (PPIs)	Esomeprazole, Lansoprazole, Omeprazole, Pantoprazole, Misoprostol[a]	£0.12	£11.24
Drugs used to treat treatment side-effects			
Proton pump inhibitors (PPIs)	Esomeprazole, Lansoprazole, Omeprazole, Pantoprazole, Rabeprazole	£0.06	[b]
Laxatives	Bisacodyl, Lactulose, Macrogol, Senna	£0.22	£3.13[c]
Treatment of vertigo	Betahistine, Cinnarazine, Cyclizine, Prochlorperazine	£0.22	£3.12[c]
Anti-emetics	Domperidone, Metoclopramide	£0.14	£1.92[c]

Abbreviations: NSAIDs= non-steroidal anti-inflammatory drugs; PPI= proton pump inhibitor

(a) Misoprostol is a prostaglandin analogue, but for the purposes of this review has been grouped with the PPIs
(b) Cost per cycle dependent on type of adverse event being treatment. For gastrointestinal bleeds, the daily cost was multiplied by 42 days' worth of treatment. For dyspepsia and symptomatic ulcer, the daily cost was multiplied by 28 days supply. The treatment dose for PPIs was assumed to be maximum dose listed in the British National Formulary.
(c) Based on a supply of 14 days treatment

2.3.5.2 Monitoring costs and osteoarthritis follow-up

It was assumed that all people with osteoarthritis attended annual follow-ups with their GP. In addition to this, patients receiving NSAIDs or opioids were also assumed to attend an additional appointment with their GP during the year, with patients receiving NSAIDs also having an annual biochemistry test. The annual cost was divided by four to give a cycle cost. The cost of a GP appointment was taken from the PSSRU 2020 [27], and the cost of a biochemistry test was taken from the NHS Reference Costs 2019/20.[74]

2.3.5.3 Adverse event costs

The costs of hospital resources (e.g., endoscopy) were taken from the NHS Reference Costs 2019/20.[74] Staff costs were taken from the PSSRU unit costs 2020.[27]

Cost of gastrointestinal bleeding

For gastrointestinal bleeding events, the methodology utilised in the previous OA guideline model was followed. Here, a decision tree was used to estimate average cost per event, based on assumptions made in the 2006 HTA paper on gastroprotection (Brown 2006).[19] Costs of each branch in the decision tree were calculated using HRG codes from the NHS reference costs 2019/20.[74] GP contacts and outpatient visits were assumed and included, again based on data from the 2006 HTA paper on gastroprotection (Brown 2006).[19] All patients with a gastrointestinal bleed were assumed to have a helicobacter test.

Table 18. Gastrointestinal bleeding resource use description

Adverse event	Resource use description
Gastrointestinal bleeding event	Two GP consultations are assumed, with a PPI prescribed for 42 days. Those patients treated in the inpatient setting are assumed to have one outpatient visit and a surgical procedure (including an endoscopy), or a therapeutic endoscopy alone if surgery is not undertaken. Patients treated in the outpatient setting are assumed to have two gastroenterology visits and a therapeutic endoscopy.

Abbreviations: GI= gastrointestinal; GP= general practitioner; PPI= proton pump inhibitor

The cost per event, along with the weighted average cost per event is reported in

Table 19.

Table 19. Cost associated with the management of gastrointestinal bleeds

Treatment pathway	Total cost
Inpatient surgical	£3,199
Inpatient medical	£946
Outpatient	£1,091
Weighted average cost of gastrointestinal bleeds	**£1,581**

Cost of non-bleeding gastrointestinal events and vertigo

For the adverse events constipation, nausea, vomiting and vertigo, a single GP consultation along with a prescription for 14 days treatment was assumed. For dyspepsia and symptomatic ulcer, a decision tree was used to estimate average cost per event, based on assumptions made in the 2006 HTA paper on gastroprotection (Brown 2006).[19]

Table 20. Dyspepsia and symptomatic ulcer resource use description

Adverse event	Resource use description
Dyspepsia	It was assumed that in the majority of cases (98%) no further investigation was undertaken after an initial consultation with the GP who prescribes a PPI for one month. A small minority of patients are treated in an inpatient setting with an endoscopy or as an outpatient with or without endoscopy.
Symptomatic ulcer	Two GP consultations are assumed who prescribes a PPI is prescribed for one month. Two gastroenterology outpatient appointments are also assumed with or without an endoscopy.
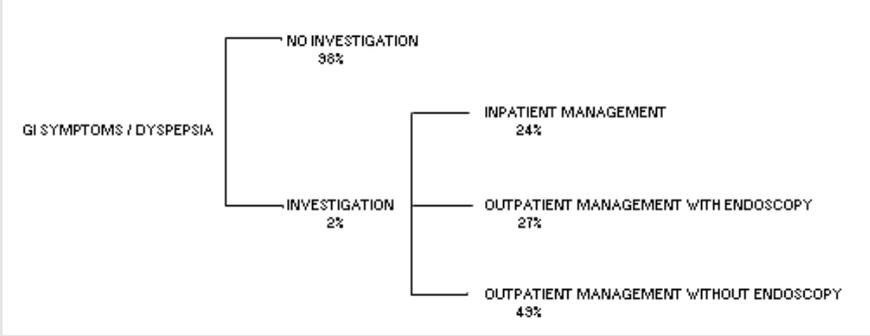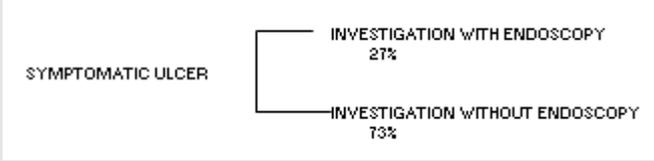
Abbreviations: GI= gastrointestinal; GP= general practitioner; PPI= proton pump inhibitor

The cost per event, along with the weighted average cost per event for dyspepsia and symptomatic ulcer are reported in Table 21 and Table 22, respectively.

Table 21. Cost associated with the management of dyspepsia

Treatment pathway	Total cost
No investigation	£41
Inpatient	£700
Outpatient endoscopy	£700
Outpatient no endoscopy	£391
Weighted average cost of gastrointestinal bleeds	**£52**

Table 22. Cost associated with the management of symptomatic ulcer

Treatment pathway	Total cost
Outpatient endoscopy	£1,090
Outpatient no endoscopy	£391
Weighted average cost of gastrointestinal bleeds	**£579**

Cost of cardiovascular events

Sources of cost data were identified by reviewing sources used in other similar cardiovascular models (NICE guideline or technology appraisal models or published

economic models) and through non-systematic online searches to identify newer publications.

Costs of stroke were based on Xu 2016[60], which undertook a patient level simulation using audit data from the UK Sentinel Stroke National Audit Programme to generate estimates of the financial burden of Stroke to the NHS and social care services. The estimates of costs attributable to stroke from resulting health and social care provision, were estimated up to five years after the first stroke. The total of 1 year and 5-year costs were reported, with NHS and social care costs being reported separately. Recurrent strokes were also included in the costs. For the event state cost in the model, the 1-year total costs from the study were used. The 1-year costs included both local authority, NHS and private social care costs, as it was not possible to disaggregate the two. The costs of the post stroke state was calculated by dividing the average cost between years 1 and 5 by the average life years between years 1 and 5, in order to derive the cost per life year. The 5-year cost included only local authority social care costs and NHS costs, as these were reported separately in the report. Since the costs of stroke and post-stroke were presented as annual sums, to convert these to 3-monthly cycle costs, the post-stroke cost was first divided by four to give a 3-monthly cost. The additional cost associated with the first year after stroke was bundled into the acute stroke state. This was achieved by taking the first-year cost of stroke and deducting the cost of three 3-monthly post-stroke cycles from it.

Danese 2016 aimed to characterise the costs to the UK National Health Service of cardiovascular (CV) events among individuals receiving lipid-modifying therapy.[29] It was a retrospective cohort study using Clinical Practice Research Datalink records from 2006 to 2012 to identify individuals with their first and second CV related hospitalisations (first event and second event cohorts). Costs were reported for TIA, unstable angina, MI, and heart failure. The study only included healthcare costs. Costs after each CV event were estimated, and the incremental difference from the period before the first CV event was calculated. The follow up period was 36 months after the event, with costs broken down into the first 6 months, and 7-36 months. The cost between 7-36 months was reported as annualised costs. This was divided by four to reflect the cost of a cycle in the chronic state under the assumption that the cost will remain constant over time. The cost between 0-6 months reflects the acute costs and assumed to occur in the first cycle of the model. The 3-monthly chronic cost was subtracted from the acute state cost to ensure the overall cost in the first six months reflects the costs reported in the source.

All published costs above were inflated to 2019/20 costs using the Hospital & Community Health Services (HCHS) Pay & Prices Index.[27]

The cost for the stable angina event state was based on a weighted average of HRG codes EB12A-D from the NHS reference costs 2019/20. The post-stable angina state was assumed to cost the same as the post-unstable angina health state.

The costs assigned to the CV health states in the model are summarised in Table 23.

Table 23: Health state costs

State	Cost (per cycle)	Source
Stroke	£12,941	Xu (2016)[86] - SSNAP project inflated to 2019/20
Post-stroke	£1,728	
TIA	£1,646	Danese (2016)[29] inflated to 2019/20
Post-TIA	£36	
Myocardial infarction	£4,427	
Post-MI	£253	
Stable angina	£687	NHS reference costs 2019/20.[74] Total HRG's. EB13. Weighted average of the complication and comorbidity codes

State	Cost (per cycle)	Source
Post-stable angina	£90	Assumed same as post unstable angina state.
Unstable angina	£2,295	Danese (2016) inflated to 2019/20[29]
Post-unstable angina	£90	
Heart failure	£2,442	
Post heart failure	£232	

Abbreviations: MI= myocardial infarction; TIA= transient ischaemic attack

Source/Note: All published costs that were inflated above were inflated to 2019/20 costs using the Hospital & Community Health Services (HCHS) Pay & Prices Index[27]

Cost of renal events

The cost of acute kidney injury was based on a weighted average of total HRG costs of AKI from the NHS reference costs 2019/20(see Table 24 below).[74]

Table 24. Cost of acute kidney injury

Acute kidney injury code	Activity	Unit cost
LA07H Acute Kidney Injury with Interventions, with CC Score 11+	2,477	£6,282
LA07J Acute Kidney Injury with Interventions, with CC Score 6-10	3,727	£4,794
LA07K Acute Kidney Injury with Interventions, with CC Score 0-5	2,765	£3,828
LA07L Acute Kidney Injury without Interventions, with CC Score 12+	9,437	£2,850
LA07M Acute Kidney Injury without Interventions, with CC Score 8-11	20,958	£2,072
LA07N Acute Kidney Injury without Interventions, with CC Score 4-7	34,676	£1,546
LA07P Acute Kidney Injury without Interventions, with CC Score 0-3	24,807	£1,045
Pooled average		**£1,961**

Cost of moderate chronic kidney disease (stages 3-4)

The cost of moderate and severe chronic kidney disease (CKD) was based on the NICE contrast-induced acute kidney injury (CI-AKI) guideline.[61] The resource use assumption underpinning that model were kept, and unit costs, staff costs and drug costs updated using the National Reference Costs 2019/20, PSSRU 2020 and Drug Tariff November 2021 edition, respectively. The original model included HRG codes associated with vascular access for renal replacement therapy for stage 5 CKD taken from the 2010/11 National Reference Costs which were no longer present in the 2019/20 version. These costs were therefore inflated to 2019/20 costs.

Costs associated with the CKD stages 3–4 include nephrology appointments, estimated glomerular filtration rate (eGFR) measurements (consisting of one biochemistry and one phlebotomy test), treatment for anaemia (with epoetin) and treatment with diuretics, (Table 25). The assumptions surrounding the resource use (e.g., the number of appointments), as well as epoetin and furosemide requirements were all based on the acute kidney kidney injury guideline (CG148).[63]

Table 25. Cost of stage 3-4 CKD per 3-month cycle

	Value	Source
Number of nephrologist appointments per cycle	1	NICE 2013a[61]
Cost of nephrology appointment	£175	NHS Reference costs 2019/20[74]
Number of eGFR measurements per cycle	1	NICE 2013a[61]
Cost of eGFR measurements	£5.42	NHS Reference costs 2019/20[74]
Proportion of people with anaemia receiving epoetin	9%	NICE 2013a[61]
Cost of epoetin per cycle	£69	Calculation (a)
Proportion of people receiving furosemide (stage 4 only)	60%	NICE 2013a[61]
Cost of furosemide per cycle	£0.42	Calculation (b)
Total cost	**£250**	Calculation (c)

Abbreviation: eGFR= estimated glomerular filtration rate
(a) Average cost of one unit of epoetin is £0.03 (Drug Tariff November 2021). Average weekly dose 1,788 units (CI-AKI guideline model). Per cycle costs obtained by multiplying the weekly dose by the cycle length (in weeks), then by the cost per unit and the percentage with anaemia.
(b) Cost of a 40 mg tablet of furosemide is £0.03 (Drug Tariff November 2021). Assumed dose is 40 mg per day (NICE, 2013a). Total cost obtained by multiplying the cost for one tablet by the number of days in a cycle, followed by the percentage receiving furosemide.
(c) Total cost of nephrology appointments, eGFR tests, epoetin and furosemide.

Cost of severe CKD (stage 5)

Patients who enter stage 5 CKD will, in addition to the drug costs outlined in

Table 25, incur costs associated with either renal replacement therapy (RRT) or conservative management (CM). They will also incur costs such as RRT access procedures, anaemia management, specialist appointments, eGFR measurements and diuretics. The costs for stage 5 CKD were calculated differently for cycle 1 and for cycle 2 onwards and are shown in Table 26.

In cycle 1, patients are initiating treatment and therefore will be receiving care with increased intensity. In this cycle, It was assumed that 90% of patients will be receiving RRT and 10% will be on CM in line with the assumptions in the CI-AKI guideline.[61]To estimate the cost of RRT, a pooled average was taken from the NHS reference costs comparing national usage of different treatment modalities with the costs per session of each modality.

Table 26: CKD Stage 5 Costs

Resource	frequency	Cost per cycle	Source of cost
Patients on RRT - Cycle 1			
Nephrologist appointment	1 per cycle	£175	NHS Reference costs 2019/20[74]
eGFR	12 per cycle	£65	NHS Reference costs 2019/20 & PSSRU 2020[27, 74]
Epoetin alpha	1788 units per week (£0.033 per unit)	£257	Drug Tariff November 2021& NICE 2013a[72]
Access procedure	1 (Haemodialysis: 79%; Peritoneal: 21%)	£1,421	NHS Reference costs 2019/20[74]
Pooled average cost of RRT (excluding APD and CAPD)	36 per cycle	£165	
Pooled average cost of APD and CAPD	84 per cycle	£77	
RRT		£6,049	Calculated[(a)]
Patients on Conservative Management (CM) – Cycle 1			
Nephrologist appointment	1 per cycle	£175	NHS Reference costs 2019/20[74]
Phone call	12 per cycle	£73	PSSRU 2020[27]
Home visits	3 per cycle	£76	
eGFR	12 per cycle	£65	NHS Reference costs 2019/20 & PSSRU 2020[27, 74]
Epoetin alpha	1788 units per week (£0.033 per unit)	£257	Drug Tariff November 2021& NICE 2013a[61, 72]
Patients on RRT cycle 2 onwards			
Nephrologist appointment (no initial consultation)	2 per cycle	£349	NHS Reference costs 2019/20[74]
eGFR	12 per cycle	£65	NHS Reference costs 2019/20 & PSSRU 2020[27, 74]
Epoetin alpha	1788 units per week (£0.033 per unit)	£39	Drug Tariff November 2021& NICE 2013a[61, 72]
Access procedure	0.15 per cycle (Haemodialysis: 79%; Peritoneal: 21%)	£213	NHS Reference costs 2010/11 (inflated to 2019/20 costs) [27, 33]
Pooled average cost (excluding APD and CAPD)	36 per cycle	£165	NHS Reference costs 2019/20[74]

Pooled average cost of APD and CAPD	84 per cycle	£77	
RRT		£6,049	Calculated[a]
Patients on CM- cycle 2 onwards			
Nephrologist appointment (no initial consultation)	2 per month	£349	NHS Reference costs 2019/20[74]
Phone call	12 per cycle	£73	PSSRU 2020[27]
Home visits	3 per cycle	£76	
eGFR	12 per cycle	£65	NHS Reference costs 2019/20 & PSSRU 2020[27, 74]
Epoetin alpha	1788 units per week (£0.005 per unit)	£257	Drug Tariff November 2021& NICE 2013a[61, 72]
Cycle 1	**90% RRT/10% CM**	**£7,236**	
Cycle 2 onwards	**90% RRT/10% CM**	**£6,323**	

Abbreviations: APD= Automated Peritoneal Dialysis; CAPD= Continuous Ambulatory Peritoneal Dialysis; CM= conservative management; eGFR= estimated glomerular filtration rate; RRT= renal replacement therapy
(a) RRT is calculated by multiplying the pooled average cost of either haemodialysis or peritoneal dialysis by the frequency per cycle and by the proportion of patients receiving haemodialysis or peritoneal dialysis

Cost of hepatic events

The cost of acute liver failure (ALF) and liver transplant were taken from the NHS Reference costs 2019/20.[74] The ICD-10 code for acute liver failure (K72.00: Acute and subacute hepatic failure without coma) was mapped to the HRG code GC01: Liver failure disorders. Given the severity of ALF, a weighted average of the HRG codes including single and multiple interventions (GC01C, GC01D) was applied to the cost of management. The cost associated with post-transplant was taken from the non-alcoholic fatty liver disease (NAFLD) model. Here, a 6-monthly cost were calculated by taking an average of the costs from patients with hepatitis' B and C. This cost was inflated to 2019/20[27], and halved to give the average 3-monthly cost. It should be noted that the NAFLD model apportioned costs to liver transplant over two years, while this model apportions the cost of a liver transplant according to the total HRG code only, therefore additional costs incurred immediately after liver transplant may not have been captured and the true cost of transplant may therefore be underestimated. The costs for related to hepatic events are presented in

Table 27.

Table 27. Costs associated with hepatic adverse events

Acute liver failure	Activity	Unit cost
GC01C Liver Failure Disorders with Multiple Interventions	810	£6,296
GC01D Liver Failure Disorders with Single Intervention	1,880	£3,519
Pooled average	**£4,355**	
Liver transplant		
GA15A Liver Transplant, 18 years and over	703	£20,827
Pooled average	**£20,827**	

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 cohorts 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.

A cohort of 1,000 people entered the Markov model for each of the interventions. People then moved between health states (defined as adverse events) based on probabilities of events occurring which was calculated using incidence data and treatment relative effects. Mortality transition probabilities in the Markov model also depend on the health states people are in. In each cycle, events are calculated based on the population at risk and used to calculate overall QALYs and costs.

Mortality rates were converted into transition probabilities for the respective cycle length (3 months in the base case) before inputting into the Markov model.

	Where
$Transition\ Probability\ (P) = 1 - e^{-rt}$	r=selected annual rate t=0.25 (the cycle length)

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 year 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:

	Where:
$Discounted\ total = \dfrac{Total}{(1+r)^n}$	r=discount rate per annum n=time (0.25 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

SA1: Lifetime treatment duration with some AEs in the first cycle only

In the base case, it is assumed that the treatment duration is 3 months. Sensitivity analyses explored the effect of a lifetime duration. Since it is likely that patients taking medication for an extended period will likely experience certain adverse events requiring withdrawal soon after treatment initiation, this scenario limited the occurrence of the following adverse events to the first cycle only: constipation, nausea, vomiting, dyspepsia, symptomatic ulcer, gastrointestinal bleeds and vertigo.

SA2: Add falls and hip fractures

Falls and hip fractures are adverse events that have a significant impact on health-related quality of life and NHS resource use. The committee acknowledged that these adverse events could be associated with drug treatment in some populations but were sceptical of their applicability as a drug-related occurrence in the osteoarthritis population, since falls and fractures were more likely to occur as a result of mechanical changes in joint movement resulting from the osteoarthritis itself. The committee agreed therefore to explore the impact of drug-related falls and hip fractures in the model during sensitivity analysis. For this analysis, data for falls and hip fractures were based on the safe prescribing of benzodiazepine guideline model.[62]

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[ref] conducted alongside the NICE guideline CG161 on falls in older people[62]as no more recent data was found. Results from the study on a subpopulation of older people (>65) are reported elsewhere.[44] The annual rates of falls are shown in Table 28.

Table 28: 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
Incidence rate ratio by age group			
65-69	1	1	
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)	HSE 2005
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%	
70-74	24%	27%	
75-79	21%	20%	HSE 2005
80-84	17%	13%	
85+	12%	8%	

Abbreviation: HSE= health survey of England
(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 61 and over, fall incidence rates for people aged between 61 to 65 had to be

extrapolated. For people aged 61 to 65 we assumed the same fall rate as the age group 65-69.

Hip fracture

The annual probability of a hip fracture was taken from a Canadian study including 21,687 fractures which occurred in people aged 50 or older as no UK data was available. Probabilities by age and gender are presented in Table 29.

Table 29: Hip fracture and annual probability

Age group	Women	Men	Source
51-60	0.0002	0.0002	
61-70	0.0008	0.0006	
71-80	0.0037	0.0020	Hopkins 2012
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.

Mortality

Data for mortality after a fall was taken from a study by Scuffham 2003[85], which looked at incidence and costs of falls in older people in the United Kingdom during 1999. The percentage of patients who died after a fall is presented below. The figures were converted to 3-monthy probabilities for the model.

Age group	Incidence of death after fall	Source
65-69	0.43%	Scuffham 2003[85]
70-74	0.65%	
75+	0.91%	

Mortality in the population with hip fracture was calculated using the relative effects shown in Table 30.

Table 30: Mortality relative effects

Event	Parameter	Value	Source
Hip fracture	Hazard ratio	Males: 7.95 (6.13 to 10.30) Females: 5.75 (4.94 to 6.69)	Haentjens 2010[43]
Post hip fracture	Hazard ratio	Males: 1.90 (1.58 to 2.3) Females: 1.86 (1.6 to 2.16)	Haentjens 2010[43]

The hazard ratios for hip fracture and post-hip fracture states were based on a meta-analysis 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 during a period ranging from 1979 to 2009. The hazard ratio for the first three months was used for the acute hip fracture health state as it 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, it was assumed that post-hip fracture state is a permanent state, where people remain until they die.

Cost of falls

The cost of a fall in calculated using a Scottish cost analysis from Craig 2013.[24] 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 31 details all fall-related events. 20% of people who fall were estimated to experience a "serious fall" by summing up those attending general practices, calling an ambulance and attending A&E.

Table 31: Falls-related events proportion

Event	Proportion	Source
Serious fall	20% of total falls	Craig 2013[24]
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 a standard UK source (NHS Reference Costs and PSSRU) and the study from Craig. All the costs used in the model and their sources are listed in Table 32.

Table 32: Falls-related events cost

Event	Cost	Source
GP visit [a]	£38	PSSRU 2020[28]
Ambulance call-out	£213	NHS Ref Costs 2019/20[74]
A&E non admitted	£155	NHS Ref Costs 2019/20[74]
A&E Admitted	£313	NHS Ref Costs 2019/20[74]
Inpatient stay (no hip fracture)	£8,195	Craig 2013[25] inflated to 2019/20 prices
Home discharge	£1,965	Craig 2013[25] inflated to 2019/20 prices
Residential: Short-term	£9,302	Craig 2013 inflated to 2019/20 prices
Residential: long-term	£72,971	Craig 2013 inflated to 2019/20 prices

(a) Including direct care and qualification costs

Based on the data presented in Table 31 and Table 32, the average cost of a fall in the UK was calculated as £1,330.

Cost of hip fracture

The cost of a hip fracture was estimated using a UK cost analysis 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.[51]

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 33. Costs included diagnostic and treatment cost both outpatient and inpatient sustained by the NHS.

Table 33: Hip fracture cost

Hip fracture cost	Mean	Standard error	Source
First year	£14,163	£254	Leal 2016[52]
Second year	£2,139	£90	

The model applies the first-year cost to those in the acute hip fracture state and the second-year cost to those in the post-hip fracture state. Since the costs of hip fracture and post-hip fracture were presented as annual sums, to convert these to 3-monthly cycle costs, the post-hip fracture cost was first divided by four to give a 3-monthly cost. The cost associated with fracture itself was bundled into the first cycle. This was achieved by taking the annual cost of hip fracture and deducting to the cost of three 3-monthly post-stroke cycles from it.

Utility decrements after falls and hip fractures

Decrements for falls and hip fractures were also taken from the NICE safe prescribing of benzodiazepine model (2021). Utility losses caused by a fall injury were identified from an existing and published model by Church 2012 reporting an annual loss of utility caused by falls-related events and are shown in Table 34.[22]

Table 34: 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[22]
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 along with their source are presented in

Table 35. 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, they 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.

Table 35: Falls-related events

Event	Proportion	Source
Serious fall	20% of total falls	Craig 2013[24]
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.

The utility decrements for the hip fracture health state and the post-hip fracture health state were also based on the benzodiazepine safe prescribing model where the post-hip fracture utility decrement was taken from a prospective cohort of 741 patients treated at a single major trauma centre in the United Kingdom.[42] The reduction in EQ-5D at 12 months was reported as -0.22. To calculate the utility decrement during the 12 months, the area above the curve EQ-5D in Figure 3 was estimated using the trapezoidal rule assuming linear change between data points.

Figure 3. Changes in EQ-5D following a hip fracture
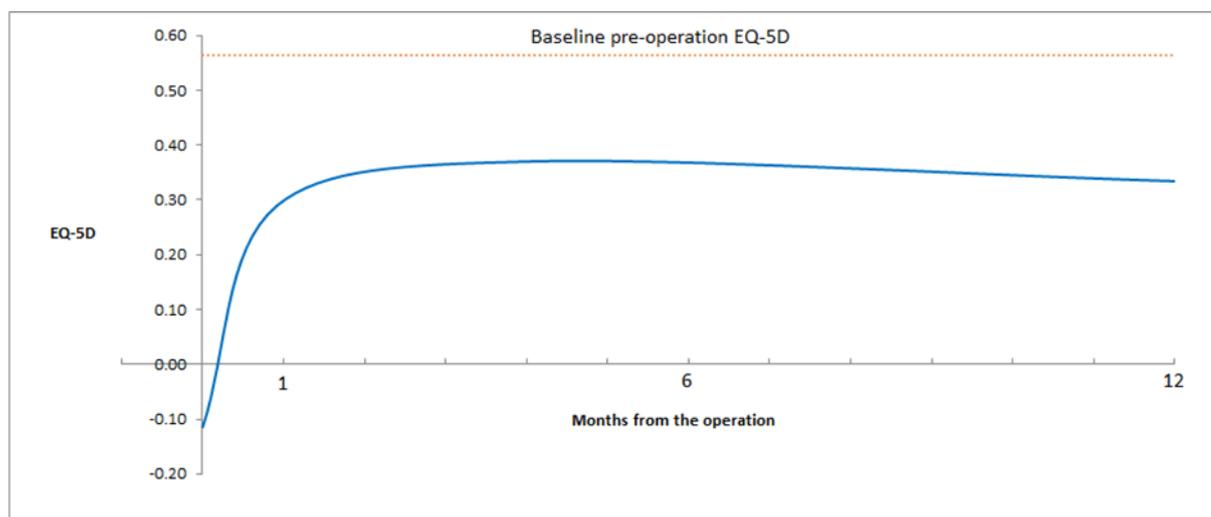
Source: Griffin 2015

This gives a QALY loss of 0.237 that was assigned to people in the acute hip fracture state.

Utility decrements for falls and hip fractures used in the mode are presented in Table 36.

Table 36. Utility decrements

Adverse event	Utility decrement
Falls injury	-0.011
Hip fracture (New event)	-0.237
Hip fracture (Post event)	-0.220

SA3: Add laxatives with opioids

The model was based on the clinical trial setting where patients were given opioids alone without preventative laxatives. However, in clinical practice, it is normative to co-prescribe a

laxative with strong opioids to mitigate the risk of opioid-related constipation. In this sensitivity analysis, the average-weighted cost of laxatives was added to the cost of opioids and in turn it was assumed that the risk of constipation in this cohort remained the same as the no treatment cohort.

SA4: Remove SA, UA, TIA

The previous osteoarthritis guideline model did not include SA, UA and TIA within CV adverse events; therefore, this analysis also excludes these to enable a better comparison of the results with previous models

SA5: NSAID+PPI cardiovascular event RR same as NSAID alone

The base case used the relative risk for CV events with NSAIDs plus PPIs that was reported in the clinical review. However, this value was based on a single trial.[26] This was substantially higher than the relative risk reported with oral NSAIDs alone versus placebo. However, a comparison between oral NSAIDs plus PPI and oral NSAIDs alone reported no significant difference between the two in relation to cardiovascular events. This scenario therefore sets the relative risk of cardiovascular events for oral NSAIDs plus PPIs to that of oral NSAIDs alone.

SA6: Use NSAIDs plus PPI RR for cardiovascular events from NMA

The base case used the relative risk for CV events with NSAIDs plus PPIs that was reported in the clinical review. However, this value was based on a single trial.[26] A sensitivity analysis explored the effect of substituting this relative risk with the relative risk derived from trials reporting adverse events with NSAIDS plus PPI versus no treatment, NSAIDs alone versus no treatment and NSAIDs plus PPIs versus NSAIDs alone (see section 2.3.1.1 for further information).

SA7: Exclude sudden cardiovascular death

In this scenario, the mortality associated with the acute phase of cardiovascular events were excluded.

SA8: Only short-term cardiovascular impact of treatment

In this scenario, only the costs associated with cardiovascular adverse events were considered.

SA9: Utilities: Barton algorithm

This analysis explored the effect of using the Barton algorithm to map between WOMAC scores and EQ-5D scores instead of the Wailoo algorithm (see section 0 for further details).

SA10: Utilities: Price algorithm

This analysis explored the effect of using the Price algorithm to map between WOMAC scores and EQ-5D scores instead of the Wailoo algorithm (see section 0 for further details).

SA11: Utilities: Lawrence algorithm

This analysis explored the effect of using the Lawrence algorithm to map between SF-36 summary scores and EQ-5D scores instead of the Price algorithm (see section 0 for further details).

SA12: Utilities: Maund algorithm

This analysis explored the effect of using the Maund algorithm to map between SF-36 summary scores and EQ-5D scores instead of the Price algorithm (see section 0 for further details).

SA13: Use drug cost from trials

The unit cost for each drug was calculated by multiplying the weighted-average cost per unit (usually milligrams) by its daily dosage. The daily dosage was taken from the online British National Formulary (BNF), which sometimes differed from what was reported in the clinical trials. For example, the recommended daily dosage of paracetamol in the BNF is 4 grams, yet the clinical trial for paracetamol used in the base case used a daily dosage of 3 grams. This sensitivity analysis applies the weighted average drug unit costs in line with dosages reported in the clinical trials, thereby applying the appropriate cost to the treatment effects.

SA14: Opioid daily dose less than or equal to 40mg morphine

There is a drug-dose to adverse event relationship observed with opioids, where higher doses result in a higher likelihood of adverse events. The committee were interested in splitting the utilities reported in clinical trials between those that reported a morphine equivalent daily dose less than or equal to 40mg and those that reported a morphine-equivalent daily dose of greater than 40mg. This threshold was chosen as it has significance in clinical practice. The conversion of various opioid doses to a morphine-equivalent dose were based on figures reported by the Faculty of Pain Medicine[35], and are presented below in Table 37.

Table 37. Equi-analgesic potencies of opioids relative to 40mg daily oral morphine

Hip fracture cost	Potency	Equivalent dose to 40mg oral morphine
Morphine (oral)	1	40mg
Oxycodone (oral)	1.5	26.4mg*
Tapentadol (oral)	0.4	100mg
Tramadol (oral)	0.1	400mg

Although the potency was reported as 1.5, the equivalent dose to 10mg morphine was reported as 6.6mg of oxycodone, which was used during calculations

This sensitivity analysis explored the effect of limiting the sourcing of utilities to only those studies with a daily morphine-equivalent dose of 40mg or less. The resulting utility scores are presented in Table 38.

Table 38. Utility scores reported in trials with MMED <=40mg

Intervention	Base case utility	Utility with MMED <=40mg
Oral opioids	0.409	0.412
Transdermal buprenorphine	0.461	0.461

SA15: Opioid daily dose greater than 40mg morphine

This sensitivity analysis explored the effect of limiting the sourcing of utilities to only those studies with a daily morphine-equivalent dose of greater than 40mg. The resulting utility scores are presented in Table 39.

Table 39. Utility scores reported in trials with MMED >40mg

Intervention	Base case utility	Utility with MMED >40mg
Oral opioids	0.409	0.406
Transdermal buprenorphine	0.461	0.461

SA16: Including all VAS trials

The base case analysis excluded studies that only reported pain VAS scores during the mapping of treatment outcomes to EQ-5D, although a mapping equation between pain VAS scores and EQ-5D exists. This was because the committee considered pain and physical function outcomes as clinically relevant outcomes, therefore it was not appropriate to map from only one outcome in the base case. Nevertheless, this sensitivity analysis included those studies reporting VAS scores, which were then combined with those studies already included in the base with a resulting overall weighted average EQ-5D score then applied to each intervention. The resulting utility scores are presented in Table 40.

Table 40. Utility scores with the addition of VAS- reported trials

Intervention	Base case utility	Utility with VAS pain included
No treatment	0.400	0.400
Paracetamol	0.435	0.409
Oral NSAIDs plus PPIs	0.458	0.450
Oral NSAIDs alone	0.458	0.430
Topical NSAIDs	0.502	0.438
Oral opioids	0.409	0.413
Transdermal buprenorphine	0.461	0.432

SA17: Paracetamol best-case scenario

A best-case scenario for paracetamol utilised the Wailoo algorithm and the drug costs from the trial only. This scenario depicts the best possible scenario for paracetamol since the Wailoo algorithm results in the highest QALY gain relative to other interventions and the cost from the trial gives the lowest cycle drug costs.

SA18: Paracetamol worst-case scenario

A worst-case scenario for paracetamol utilised the Barton algorithm and included all VAS trials, since this was associated with the lowest QALY gain relative to other interventions. The cost of paracetamol based on BNF-recommended doses was also used since this represented the highest possible drug cost per cycle.

SA19: Oral NSAIDs plus PPI best case

A scenario that combined all the separate sensitivity analyses favourable towards oral NSAIDs plus PPIs was run to indicate the best possible cost per QALY. In this scenario, the relative risk of cardiovascular events was set to that of oral NSAIDs alone, the adverse events SA, US and TIA were removed and only the acute phase cost of CV events included.

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 National Guideline Centre; this included systematic checking 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

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

$$Net\,Health\,Benefit\,(X) = (QALYs(X)) - Costs(X)/\lambda$$

Cost effective if:
- Highest net benefit

Where: λ = threshold (£20,000 per QALY gained)

For ease of computation NHB is used in this analysis to identify the optimal strategy.

2.8 Interpreting results

NICE sets out the principles that committees should consider when judging whether an intervention offers good value for money. [66, 68, 71] 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 NHB to rank the strategies on the basis of their relative cost effectiveness. The the optimal strategy at a willingness to pay of £20,000 per QALY gained is the one that has the highest NHB.

3 Results

3.1 Base case

The probabilistic results of the base case scenario are presented in Table 43. They show that all interventions are cost effective compared with no treatment except for oral NSAIDs plus PPI (cost per QALY gained of £28,190) and strong oral opioids (cost per QALY gained of £32,916).

A comparison of probabilistic incremental net health benefit shows that there were four treatments ranked higher than no treatment. Topical NSAIDs are the optimal treatment strategy, followed by oral NSAIDs alone, followed by paracetamol and lastly buprenorphine.

In the sensitivity analyses topical NSAIDs never fell below a ranking of second.

Oral NSAIDs remained second in the base case and most sensitivity analyses and was never ranked below no treatment.

Paracetamol, meanwhile, was ranked third in the base case and in most sensitivity analyses. It ranked below no treatment when VAS trials were included in the analysis.

Buprenorphine ranged between third and sixth and only fell below no treatment in the rankings when falls and hip fractures were included in the analysis.

Oral NSAIDs plus PPI were generally ranked below no treatment in the rankings. However, it ranked higher than no treatment when the Barton and Price algorithms were used instead of the Wailoo algorithm and when certain assumptions regarding adverse events were relaxed (the removal of SA, UA and TIA from CV adverse events, where the relative risk of CV events was the same as NSAIDs alone, when acute mortality associated with CV events were excluded and where only the short-term cost of CV events were included.

Lastly, strong opioids ranged between third and seventh and only ranked higher than no treatment when the Maund algorithm was used, where VAS trials were included and where the morphine-equivalent daily dose was less than or equal to 40mg.

Table 41. Events per 1,000 patients (probabilistic base case results)

	No treatment	Paracetamol	Oral NSAIDs + PPIs	Oral NSAIDs alone	Topical NSAIDs	Strong opioids	Transdermal buprenorphine
GI non-bleeds	187	187	187	188	187	299	415
GI bleeds	1	1	2	2	1	1	1
Cardiovascular	441	441	443	441	441	441	441
Hepatorenal	42	42	42	42	42	41	41
Central nervous system	82	82	82	82	82	133	187

Note: GI non-bleeds, GI bleeds and central nervous system adverse events were measured in the first cycle only. Cardiovascular and hepatorenal events were measured over the model time horizon.

Table 42. Costs (probabilistic base case results)

	No treatment	Paracetamol	Oral NSAIDs + PPIs	Oral NSAIDs alone	Topical NSAIDs	Strong opioids	Transdermal buprenorphine
Drug	£0	£24	£28	£17	£28	£51	£82
Drug monitoring	£0	£0	£10	£10	£10	£10	£10
OA follow up	£594	£594	£593	£594	£594	£594	£594
GI non-bleeds	£10	£10	£10	£10	£10	£14	£19
GI bleeds	£1	£1	£3	£4	£1	£1	£1
Cardiovascular	£5,185	£5,185	£5,259	£5,191	£5,185	£5,178	£5,170
Hepatorenal	£73	£73	£76	£75	£73	£73	£73
Central nervous system	£3	£3	£3	£3	£3	£6	£8
Total costs	£5,867	£5,891	£5,983	£5,904	£5,906	£5,927	£5,956

Table 43. Cost effectiveness (probabilistic base case results)

	No treatment	Paracetamol	Oral NSAIDs + PPIs	Oral NSAIDs alone	Topical NSAIDs	Strong opioids	Transdermal buprenorphine
Total costs	£5,867	£5,891	£5,983	£5,904	£5,906	£5,927	£5,956
Life years (undiscounted)	22.82	22.82	22.79	22.81	22.82	22.82	22.82
QALYs	5.5683	5.5756	5.5724	5.5791	5.5818	5.5701	5.5777
Incr. cost (vs no treatment)	£0	£24	£116	£37	£38	£60	£89
Incr. QALYs (vs no treatment)	-	0.0073	0.0041	0.0108	0.0135	0.0018	0.0094
ICER (n versus no treatment	-	£3,301	£28,190	£3,449	£2,847	£32,916	£9,454
NHB @20k threshold	5.27	5.28	5.27	5.28	5.29	5.27	5.28
Rank of NHB	5	3	7	2	1	6	4

Figure 4. Cost effectiveness plane (probabilistic base case results)
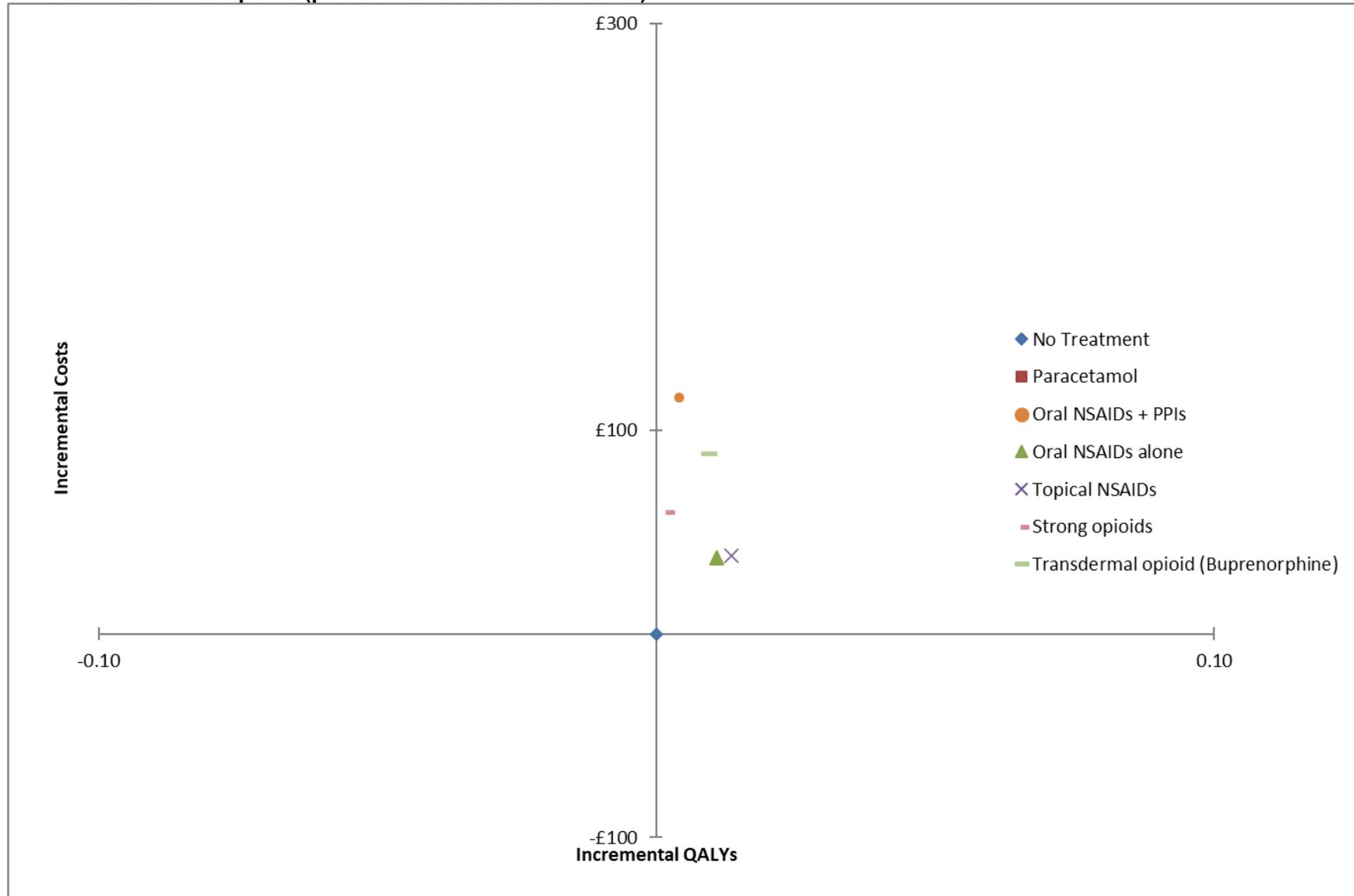
3.2 Sensitivity analyses

Table 44. Rank of net health benefit (£20,000 per QALY gained) (sensitivity analyses)

Analysis	Costs						
	No treatment	Paracetamol	Oral NSAIDs + PPIs	Oral NSAIDs alone	Topical NSAIDs	Strong opioids	Buprenorphine
Base case results	5	3	6	2	1	7	4
Treatment duration							
Lifetime treatment duration with some AEs in the first cycle only	7	3	6	2	1	5	4
Adverse events							
Add falls and hip fractures	5	3	4	2	1	7	6
Add laxatives with opioids	6	3	5	2	1	7	4
Remove SA, UA, TIA	6	3	4	2	1	7	5
NSAID+PPI CV event RR same as NSAID alone	6	4	3	2	1	7	5
Use NSAIDs plus PPI RR for CV events from NMA	5	3	7	2	1	6	4
Exclude sudden CV deaths	6	3	5	2	1	7	4
Only short-term CV impact of treatment	6	3	4	2	1	7	5
Utilities							
Utilities: Barton algorithm	6	3	5	2	1	7	4
Utilities: Price algorithm	6	3	5	2	1	7	4
Utilities: Lawrence algorithm	5	3	6	2	1	7	4
Utilities: Maund algorithm	6	2	7	4	1	5	3
Costs							
Use drug cost from trials	5	3	7	2	1	6	4
Other							
Including all VAS trials	5	7	6	2	1	4	3

Strong opioids =<40mg MMED trials	6	3	7	2	1	5	4
Strong opioids >40mg MMED trials	5	3	6	2	1	7	4
Paracetamol best case scenario	5	3	7	2	1	6	4
Paracetamol worst case scenario	5	7	6	2	1	4	3
Oral NSAIDs plus PPI best case	6	4	3	2	1	7	5

4 Discussion

4.1 Summary of results

In a comparison of paracetamol and no treatment

- One cost utility analysis reported that paracetamol was cost effective and no treatment (£12,771), however, in a full incremental analysis it was extendedly dominated by oral NSAIDs plus PPIs. This analysis was assessed as directly applicable with potentially serious limitations.
- One cost utility analysis reported that paracetamol dominated by no treatment. This analysis was assessed as directly applicable with minor limitations.
- One original cost utility analysis from this guideline review reported that paracetamol was cost effective compared with no treatment (£3,225). This analysis was assessed as directly applicable with minor limitations.

In a comparison of oral NSAIDs alone and no treatment

- One cost utility analysis reported that diclofenac, ibuprofen, naproxen and COX-2 inhibitors celecoxib and etoricoxib were dominated by no treatment. This analysis was assessed as directly applicable with potentially serious limitations
- One cost utility analysis reported that diclofenac, ibuprofen and naproxen were dominated by no treatment. However, COX-2 inhibitors celecoxib and etoricoxib were cost effective. This analysis was assessed as directly applicable with minor limitations.
- One original cost utility analysis from this guideline review reported that oral NSAIDs alone were cost effective compared with no treatment (£3,402). This analysis was assessed as directly applicable with minor limitations.

In a comparison of oral NSAIDs plus PPIs and no treatment

- One cost utility analysis reported that oral NSAIDs plus PPIs were cost effective compared with no treatment. This analysis was assessed as directly applicable with potentially serious limitations
- One cost utility analysis reported that oral NSAIDs plus PPIs were cost effective compared with no treatment. This analysis was assessed as directly applicable with minor limitations.
- One original cost utility analysis from this guideline review reported that oral NSAIDs plus PPIs were not cost effective compared with no treatment (£21,543). This analysis was assessed as directly applicable with minor limitations.

In a comparison of oral NSAIDs plus PPIs and oral NSAIDs alone

- One cost utility analysis reported that meloxicam alone was the most cost-effective strategy in a full incremental analysis. It was assessed as directly applicable with potentially serious limitations.
- One cost utility analysis reported that celecoxib 200mg plus PPI was the most cost-effective strategy in a full incremental analysis (£12,557). It was assessed as directly applicable with potentially serious limitations.
- One cost utility analysis reported that etoricoxib 200mg plus PPI was the most cost-effective strategy in a full incremental analysis (13,160). Fixed-dose combinations (ketoprofen 200mg/omeprazole 20mg, diclofenac 150mg/misoprostol 400mg and naproxen 1000mg/esomeprazole 40mg) were also dominated by NSAIDs plus PPI. This analysis was assessed as directly applicable with minor limitations.

- One original cost utility analysis from this guideline review reported that oral NSAIDs plus PPIs were dominated by oral NSAIDs alone. This analysis was assessed as directly applicable with minor limitations.

In a comparison of topical NSAIDs and no treatment

- One original cost utility analysis from this guideline review reported that topical NSAIDs were cost effective compared with no treatment (£2,744). This analysis was assessed as directly applicable with minor limitations.

In a comparison of oral NSAIDs and topical NSAIDs

- One cost utility analysis from this guideline review reported that oral ibuprofen was cost effective compared with topical ibuprofen (£9,114). This analysis was assessed as partially applicable with potentially serious limitations.
- One original cost utility analysis reported that oral NSAIDs with and without PPIs were dominated by topical NSAIDs. This analysis was assessed as directly applicable with minor limitations.

In a comparison of strong oral opioids and no treatment

- One original cost utility analysis from this guideline review reported that strong oral opioids were not cost effective compared with no treatment (£27,765). This analysis was assessed as directly applicable with minor limitations.

In a comparison of transdermal buprenorphine and no treatment

- One original cost utility analysis from this guideline review reported that transdermal buprenorphine was cost effective compared with no treatment (£8,020). This analysis was assessed as directly applicable with minor limitations.

In a comparison of glucosamine with no treatment or paracetamol

- One cost utility analysis reported that glucosamine was not cost effective compared with no treatment (£21,335). This analysis was assessed as directly applicable with minor limitations.
- One cost utility reported thar prescription crystalline glucosamine sulphate was cost effective compared to no treatment (£10,203), but that other forms of glucosamine were dominated by no treatment. This analysis was assessed as partially applicable with potentially serious limitations.
- One cost utility analysis reported that glucosamine was cost effective compared with no treatment (£3,488). It also reported that that glucosamine dominated paracetamol. This analysis was assessed as partially applicable with potentially serious limitations.

4.2   Limitations and interpretation

There were a few noteworthy limitations with this analysis. Firstly, trials that informed the treatment effects were conducted in different populations. For example, patients in trials assessing the efficacy opioids may have already tried NSAIDs and were moving up the treatment of pain ladder or may have been unable to take NSAIDs due to co-morbidities. This makes a direct comparison between different interventions difficult as it is unlikely to be a like-for-like comparison.

Secondly, this was an analysis of drug classes rather than individual drugs. As a result, the treatments in trials were heterogenous. The results of the trials used in the analysis were therefore sensitive to the frequency of particular drugs used in trials. For simplification, it was assumed that the efficacy and adverse event profile of individual drugs were broadly

representative of the overall drug class, however, in reality this is not always the case. For example, there is a higher risk of cardiovascular events with diclofenac than, say, with naproxen.

Thirdly, EQ-5D-3L scores were usually not available directly and therefore were mapped from other health outcomes (i.e., SF-36 domain and summary scores and WOMAC scores). The disadvantage of such an approach is the increased uncertainty associated with the final EQ-5D scores.

In addition, the adverse events reported in trials were mainly minor events, but in some cases they have been applied to major adverse events. For example, the main hepatic adverse event reported was elevated liver function tests, but the relative risk was applied to acute liver failure. However, since the incidence rate for liver failure was so low to begin with, a higher relative risk is unlikely to have much of an impact on the final results.

Finally, it should be noted that adverse event costs and utilities were gathered from various sources and may not always have been the most up-to-date.

4.3 Generalisability to other populations or settings

The model takes a UK perspective and may not be applicable to other settings. The average age of people in the model was 61 years, and the outcomes may not be relevant or applicable to people below this age.

It is also important to consider the location of osteoarthritis when assessing generalisability of model results since the majority of the evidence for oral treatments were based on people with knee or hip osteoarthritis, while the data for topical NSAIDs was based on people with knee or hand osteoarthritis.

It should also be noted that surgery was not included in this model, and therefore results are reflective of people who cannot or do not wish to have surgery. It is expected that surgery will be the most cost-effective option in other people.

Finally, exercise is considered the first-line treatment, and although it is not modelled here, it is expected that this model reflects people for whom exercise is not a sufficient intervention to manage the symptoms of OA.

4.4 Comparisons with published studies

There are no published studies that compared drug classes in osteoarthritis; therefore, there is a lack of clear reference point for this analysis. During the health economic review, three studies were identified in a UK setting that evaluated the cost effectiveness of individual oral NSAIDs, two of these studies also included paracetamol.

Chan 2009 included oral NSAIDs alone and with PPIs and reported that meloxicam was the most cost effective NSAID at a cost per QALY gained threshold of £20,000.

The other two studies (CG59 and CG177) reported that the combination of a COX-2 inhibitor plus PPI was the most cost effective strategy. In CG177, paracetamol was dominated by no treatment. This differs from what is reported here because the adverse events profile for paracetamol in CG177 differed.

There were no economic evaluations identified in a UK setting that compared opioids and topical NSAIDs to no treatment.

4.5   Conclusions

Given the assumptions used in the model and current drug costs, topical NSAIDs are the most cost-effective treatment for people with osteoarthritis aged 60 years and over. Other cost effective treatment options compared to no treatment included oral NSAIDs, buprenorphine and paracetamol. Oral NSAIDs plus PPIs and strong oral opioids were not cost effective in the base case analysis but were cost effective in some scenarios.

4.6   Implications for future research

This is thought to be the first model to evaluate the cost-effectiveness of drug classes in the management of osteoarthritis.

Given that the committee has recommended the withdrawal for paracetamol as an option for the management of osteoarthritis, it is important to consider all potential alternatives, including weak oral opioids. However, weak oral opioids were not included in the model due to a lack of adverse events data compared to placebo. This is the most important question not adequately answered via a health economic analysis and future research in this area is warranted.

References

1.   Afilalo M, Etropolski MS, Kuperwasser B, Kelly K, Okamoto A, Van Hove I et al. Efficacy and safety of Tapentadol extended release compared with oxycodone controlled release for the management of moderate to severe chronic pain related to osteoarthritis of the knee: a randomized, double-blind, placebo- and active-controlled phase III study. Clinical Drug Investigation. 2010; 30(8):489-505

2.   Ali T, Khan I, Simpson W, Prescott G, Townend J, Smith W et al. Incidence and outcomes in acute kidney injury: a comprehensive population-based study. Journal of the American Society of Nephrology. 2007; 18(4):1292-1298

3.   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

4.   Arshad A, Ayaz A, Rehman S, Ukrani RD, Akbar I, Jamil B. Progression of Acute Kidney Injury to Chronic Kidney Disease in Sepsis Survivors: 1-Year Follow-Up Study. Journal of Intensive Care Medicine. 2021; 36(11):1366-1370

5.   Baer PA, Thomas LM, Shainhouse Z. Treatment of osteoarthritis of the knee with a topical diclofenac solution: a randomised controlled, 6-week trial [ISRCTN53366886]. BMC Musculoskeletal Disorders. 2005; 6:44

6.   Baraf HS, Gold MS, Clark MB, Altman RD. Safety and efficacy of topical diclofenac sodium 1% gel in knee osteoarthritis: a randomized controlled trial. Physician & Sportsmedicine. 2010; 38(2):19-28

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

8.   Barthel HR, Haselwood D, Longley S, 3rd, Gold MS, Altman RD. Randomized controlled trial of diclofenac sodium gel in knee osteoarthritis. Seminars in Arthritis and Rheumatism. 2009; 39(3):203-212

9.   Barton GR, Sach TH, Jenkinson C, Avery AJ, Doherty M, Muir KR. Do estimates of cost-utility based on the EQ-5D differ from those based on the mapping of utility scores? Health Qual Life Outcomes. 2008; 6:51

10.   Belsey J, Greenfield S, Candy D, Geraint M. Systematic review: impact of constipation on quality of life in adults and children. Alimentary Pharmacology and Therapeutics. 2010; 31(9):938-949

11.   Bensen WG, Fiechtner JJ, McMillen JI, Zhao WW, Yu SS, Woods EM et al. Treatment of osteoarthritis with celecoxib, a cyclooxygenase-2 inhibitor: a randomized controlled trial. Mayo Clinic Proceedings. 1999; 74(11):1095-1105

12.   Bernal W, Hyyrylainen A, Gera A, Audimoolam VK, McPhail MJ, Auzinger G et al. Lessons from look-back in acute liver failure? A single centre experience of 3300 patients. Journal of Hepatology. 2013; 59(1):74-80

13.   Bhatia A, Goni V, Chopra S, Singh B, Katare OP. Evaluation of efficacy and safety of a novel lipogel containing diclofenac: A randomized, placebo controlled, double-blind clinical trial in patients with signs and symptoms of osteoarthritis. Contemporary Clinical Trials Communications. 2020; 20:100664

14. Bocanegra TS, Weaver AL, Tindall EA, Sikes DH, Ball JA, Wallemark CB et al. Diclofenac/misoprostol compared with diclofenac in the treatment of osteoarthritis of the knee or hip: a randomized, placebo controlled trial. Arthrotec Osteoarthritis Study Group. Journal of Rheumatology. 1998; 25(8):1602-1611

15. Bolten W, Gomes JA, Stead H, Geis GS. The gastroduodenal safety and efficacy of the fixed combination of diclofenac and misoprostol in the treatment of osteoarthritis. British Journal of Rheumatology. 1992; 31(11):753-758

16. Bookman AA, Williams KS, Shainhouse JZ. Effect of a topical diclofenac solution for relieving symptoms of primary osteoarthritis of the knee: a randomized controlled trial. CMAJ Canadian Medical Association Journal. 2004; 171(4):333-338

17. Bronnum-Hansen H, Davidsen M, Thorvaldsen P, Danish MSG. Long-term survival and causes of death after stroke. Stroke. 2001; 32(9):2131-2136

18. Bronnum-Hansen H, Jorgensen T, Davidsen M, Madsen M, Osler M, Gerdes LU et al. Survival and cause of death after myocardial infarction: the Danish MONICA study. Journal of Clinical Epidemiology. 2001; 54(12):1244-1250

19. Brown TJ, Hooper L, Elliott RA, Payne K, Webb R, Roberts C et al. A comparison of the cost-effectiveness of five strategies for the prevention of non-steroidal anti-inflammatory drug-induced gastrointestinal toxicity: a systematic review with economic modelling. Health Technology Assessment. 2006; 10(38):iii-iv, xi-xiii, 1-183

20. Chan KKW, Willan AR, Gupta M, Pullenayegum E. Underestimation of uncertainties in health utilities derived from mapping algorithms involving health-related quality-of-life measures: statistical explanations and potential remedies. Medical Decision Making. 2014; 34(7):863-872

21. Chan WP, Hsu SM, Huang GS, Yao MS, Chang YC, Ho WP. Creation of a reflecting formula to determine a patient's indication for undergoing total knee arthroplasty. Journal of Orthopaedic Science. 2010; 15(1):44-50

22. 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

23. Cowie MR, Wood DA, Coats AJ, Thompson SG, Poole-Wilson PA, Suresh V et al. Incidence and aetiology of heart failure; a population-based study. European Heart Journal. 1999; 20(6):421-428

24. 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

25. 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

26. Cryer BL, Sostek MB, Fort JG, Svensson O, Hwang C, Hochberg MC. A fixed-dose combination of naproxen and esomeprazole magnesium has comparable upper gastrointestinal tolerability to celecoxib in patients with osteoarthritis of the knee: results from two randomized, parallel-group, placebo-controlled trials. Annals of Medicine. 2011; 43(8):594-605

27. Curtis L, Burns A. Unit costs of health and social care 2020. Canterbury. University of Kent, 2020. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2020/

28.   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/

29.   Danese MD, Gleeson M, Kutikova L, Griffiths RI, Azough A, Khunti K et al. Estimating the economic burden of cardiovascular events in patients receiving lipid-modifying therapy in the UK. BMJ Open. 2016; 6(8):e011805

30.   Dawwas MF, Gimson AE, Lewsey JD, Copley LP, van der Meulen JH. Survival after liver transplantation in the United Kingdom and Ireland compared with the United States. Gut. 2007; 56(11):1606-1613

31.   DeLemos BP, Xiang J, Benson C, Gana TJ, Pascual ML, Rosanna R et al. Tramadol hydrochloride extended-release once-daily in the treatment of osteoarthritis of the knee and/or hip: a double-blind, randomized, dose-ranging trial. American Journal of Therapeutics. 2011; 18(3):216-226

32.   Dennis M, Bamford J, Sandercock P, Warlow C. Prognosis of transient ischemic attacks in the Oxfordshire Community Stroke Project. Stroke. 1990; 21(6):848-853

33.   Department of Health. **{Confirm access date with HE}** NHS reference costs 2010-11. 2012. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140 Last accessed:

34.   Eriksen BO, Ingebretsen OC. The progression of chronic kidney disease: a 10-year population-based study of the effects of gender and age. Kidney International. 2006; 69(2):375-382

35.   Faculty of pain medicine of the Royal College of A. Dose equivalents adn changing opioids. 2022. Available from: https://fpm.ac.uk/opioids-aware-structured-approach-opioid-prescribing/dose-equivalents-and-changing-opioids Last accessed: 22/02/2022.

36.   Gana TJ, Pascual ML, Fleming RR, Schein JR, Janagap CC, Xiang J et al. Extended-release tramadol in the treatment of osteoarthritis: a multicenter, randomized, double-blind, placebo-controlled clinical trial. Current Medical Research and Opinion. 2006; 22(7):1391-1401

37.   Ghosh S, Paul S, Das N, Bhattacharyya TK. A study on the effects of diclofenac sodium and etoricoxib in the treatment of osteoarthritis. Journal of the Indian Medical Association. 2007; 105(5):260-262

38.   Giansiracusa JE, Donaldson MS, Koonce ML, Lefton TE, Ruoff GE, Brooks CD. Ibuprofen in osteoarthritis. Southern Medical Journal. 1977; 70(1):49-52

39.   Golden HE, Moskowitz RW, Minic M. Analgesic efficacy and safety of nonprescription doses of naproxen sodium compared with acetaminophen in the treatment of osteoarthritis of the knee. American Journal of Therapeutics. 2004; 11(2):85-94

40.   Goldstein JL, Cryer B, Amer F, Hunt B. Celecoxib plus aspirin versus naproxen and lansoprazole plus aspirin: a randomized, double-blind, endoscopic trial. Clinical Gastroenterology and Hepatology. 2007; 5(10):1167-1174

41.   Gottesdiener K, Schnitzer T, Fisher C, Bockow B, Markenson J, Ko A et al. Results of a randomized, dose-ranging trial of etoricoxib in patients with osteoarthritis. Rheumatology. 2002; 41(9):1052-1061

42.   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. The bone & joint journal. 2015; 97(3):372-382

43.   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

44.   Health, Social Care Information C. Health survey for England - 2005, health of older people. 2009;

45.   Herrero-Beaumont G, Ivorra JA, Del Carmen Trabado M, Blanco FJ, Benito P, Martin-Mola E et al. Glucosamine sulfate in the treatment of knee osteoarthritis symptoms: a randomized, double-blind, placebo-controlled study using acetaminophen as a side comparator. Arthritis and Rheumatism. 2007; 56(2):555-567

46.   James MT, Hemmelgarn BR, Wiebe N, Pannu N, Manns BJ, Klarenbach SW et al. Glomerular filtration rate, proteinuria, and the incidence and consequences of acute kidney injury: a cohort study. Lancet. 2010; 376(9758):2096-2103

47.   Kiadaliri AA, Englund M. Assessing the external validity of algorithms to estimate EQ-5D-3L from the WOMAC. Health Qual Life Outcomes. 2016; 14(1):141

48.   Kivitz A, Eisen G, Zhao WW, Bevirt T, Recker DP. Randomized placebo-controlled trial comparing efficacy and safety of valdecoxib with naproxen in patients with osteoarthritis. Journal of Family Practice. 2002; 51(6):530-537

49.   Laine L, Harper S, Simon T, Bath R, Johanson J, Schwartz H et al. A randomized trial comparing the effect of rofecoxib, a cyclooxygenase 2-specific inhibitor, with that of ibuprofen on the gastroduodenal mucosa of patients with osteoarthritis. Rofecoxib Osteoarthritis Endoscopy Study Group. Gastroenterology. 1999; 117(4):776-783

50.   Lawrence WF, Fleishman JA. Predicting EuroQoL EQ-5D preference scores from the SF-12 Health Survey in a nationally representative sample. Medical Decision Making. 2004; 24(2):160-169

51.   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

52.   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

53.   Leung AT, Malmstrom K, Gallacher AE, Sarembock B, Poor G, Beaulieu A et al. Efficacy and tolerability profile of etoricoxib in patients with osteoarthritis: A randomized, double-blind, placebo and active-comparator controlled 12-week efficacy trial. Current Medical Research and Opinion. 2002; 18(2):49-58

54.   Lohmander LS, McKeith D, Svensson O, Malmenas M, Bolin L, Kalla A et al. A randomised, placebo controlled, comparative trial of the gastrointestinal safety and efficacy of AZD3582 versus naproxen in osteoarthritis. Annals of the Rheumatic Diseases. 2005; 64(3):449-456

55.   MacPherson H, Vickers A, Bland M, Torgerson D, Corbett M, Spackman E et al. Acupuncture for chronic pain and depression in primary care: a programme of research. Acupuncture for chronic pain and depression in primary care: a programme of research. Programme Grants for Applied Research. Southampton (UK). 2017.

56.   Maetzel A, Krahn M, Naglie G. The cost effectiveness of rofecoxib and celecoxib in patients with osteoarthritis or rheumatoid arthritis. Arthritis Care and Research. 2003; 49(3):283-292

57.   Maund E, Craig D, Suekarran S, Neilson A, Wright K, Brealey S et al. Management of frozen shoulder: a systematic review and cost-effectiveness analysis. Health Technology Assessment. 2012; 16(11):1-264

58.   Melo Gomes JA, Roth SH, Zeeh J, Bruyn GA, Woods EM, Geis GS. Double-blind comparison of efficacy and gastroduodenal safety of diclofenac/misoprostol, piroxicam, and naproxen in the treatment of osteoarthritis. Annals of the Rheumatic Diseases. 1993; 52(12):881-885

59.   National Clinical Guideline Centre L. Unstable angina and NSTEMI: the early management of unstable angina and non-ST-segment-elevation myocardial infarction. NICE clinical guideline 94. 2009;

60.   National Guideline C. Sentinel Stroke National Audit Programme. Sentinel Stroke National Audit Programme: Cost and cost-effectiveness analysis. 2016;

61.   National Institute for H, Clinical E. Acute kidney injury: prevention, detection and management [CG169]. 2013;

62.   National Institute for H, Clinical E. Falls in older people: assessing risk and prevention. Clinical guideline CG161. Appendix K: Full health economic report. 2013;

63.   National Institute for Health and Care Excellence. Acute kidney injury: prevention, detection and management. NICE clinical guideline 148. London. National institute for Health and Care Excellence, 2019. Available from: https://www.nice.org.uk/guidance/ng148

64.   National institute for Health and Care Excellence. British National Formulary. Available from: https://bnf.nice.org.uk/ Last accessed: 10/02/2022.

65.   National Institute for Health and Care Excellence. Cardiovascular disease: risk assessment and reduction, including lipid modification: Clinical Guideline 181. London. 2016. Available from: https://www.nice.org.uk/guidance/cg181

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

67.   National Institute for Health and Care Excellence. Hypertension in adults: diagnosis and management. NICE guideline 136. London. National Institute for Health and Care Excellence, 2019. Available from: https://www.nice.org.uk/guidance/ng136

68.   National Institute for Health and Care Excellence. The NICE Charter. 2020. Available from: https://www.nice.org.uk/about/who-we-are/our-charter Last accessed: 10/02/2022.

69.   National Institute for Health and Care Excellence. Non-alcoholic fatty liver disease (NAFLD): assessment and management. NICE guideline 49. London. National Institute for Health and Care Excellence, 2016. Available from: https://www.nice.org.uk/guidance/ng49

70.   National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal 2013. London. National Institute for Health and Clinical Excellence, 2013. Available from: http://publications.nice.org.uk/pmg9

71.   National Institute for Health and Clinical Excellence. Social value judgements: principles for the development of NICE guidance. London. National Institute for Health and Clinical Excellence, 2008. Available from: https://www.nice.org.uk/media/default/about/what-we-do/research-and-development/social-value-judgements-principles-for-the-development-of-nice-guidance.pdf

72.   NHS Business Services Authority. NHS electronic drug tariff November 2021. 2022. Available from: https://www.nhsbsa.nhs.uk/pharmacies-gp-practices-and-appliance-contractors/drug-tariff Last accessed: 10/02/2022.

73.   NHS Business Services Authority. Prescription Cost Analysis data. 2022. Available from: https://www.nhsbsa.nhs.uk/prescription-data/dispensing-data/prescription-cost-analysis-pca-data Last accessed: 10/02/2022.

74.   NHS England and NHS Improvement. National Cost Collection Data Publication 2019-2020. London. 2020. Available from: https://www.england.nhs.uk/wp-content/uploads/2021/06/National-Cost-Collection-2019-20-Report-FINAL.pdf

75.   Office for National Statistics. National life tables, UK: National Life Tables. London. 2021. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables

76.   Price A, Smith J, Dakin H, Kang S, Eibich P, Cook J et al. The Arthroplasty Candidacy Help Engine tool to select candidates for hip and knee replacement surgery: development and economic modelling. Health Technology Assessment. 2019; 23(32):1-216

77.   Roberts SE, Button LA, Williams JG. Prognosis following upper gastrointestinal bleeding. PloS One. 2012; 7(12):e49507

78.   Rosengren A, Wilhelmsen L, Hagman M, Wedel H. Natural history of myocardial infarction and angina pectoris in a general population sample of middle-aged men: a 16-year follow-up of the Primary Prevention Study, Goteborg, Sweden. Journal of Internal Medicine. 1998; 244(6):495-505

79.   Roth SH, Shainhouse JZ. Efficacy and safety of a topical diclofenac solution (pennsaid) in the treatment of primary osteoarthritis of the knee: a randomized, double-blind, vehicle-controlled clinical trial. Archives of Internal Medicine. 2004; 164(18):2017-2023

80.   Schmitt W, Walter K, Kurth HJ. Clinical trial on the efficacy and safety of different diclofenac formulations: multiple-unit formulations compared to enteric coated tablets in patients with activated osteoarthritis. Inflammopharmacology. 1999; 7(4):363-375

81.   Schnitzer TJ, Dattani ID, Seriolo B, Schneider H, Moore A, Tseng L et al. A 13-week, multicenter, randomized, double-blind study of lumiracoxib in hip osteoarthritis. Clinical Rheumatology. 2011; 30(11):1433-1446

82.   Schnitzer TJ, Hochberg MC, Marrero CE, Duquesroix B, Frayssinet H, Beekman M. Efficacy and safety of naproxcinod in patients with osteoarthritis of the knee: a 53-week prospective randomized multicenter study. Seminars in Arthritis and Rheumatism. 2011; 40(4):285-297

83.   Schubiger BI, Ciccolunghi SN, Tanner K. Once daily dose treatment with a non-steroidal anti-rheumatic drug (diclofenac) in osteoarthrosis. Journal of International Medical Research. 1980; 8(2):167-174

84.   Scott DL, Berry H, Capell H, Coppock J, Daymond T, Doyle DV et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. Rheumatology. 2000; 39(10):1095-1101

85.   Scuffham P, Chaplin S, Legood R. Incidence and costs of unintentional falls in older people in the United Kingdom. Journal of Epidemiology and Community Health. 2003; 57(9):740-744

86.   Sentinel Stroke National Audit P. Sentinel Stroke National Audit Programme: Cost and cost-effectiveness analysis. 2016;

87.   Serrie A, Lange B, Steup A. Tapentadol prolonged-release for moderate-to-severe chronic osteoarthritis knee pain: a double-blind, randomized, placebo- and oxycodone controlled release-controlled study. Current Medical Research and Opinion. 2017; 33(8):1423-1432

88.   Sikes DH, Agrawal NM, Zhao WW, Kent JD, Recker DP, Verburg KM. Incidence of gastroduodenal ulcers associated with valdecoxib compared with that of ibuprofen and diclofenac in patients with osteoarthritis. European Journal of Gastroenterology and Hepatology. 2002; 14(10):1101-1111

89.   Simon LS, Grierson LM, Naseer Z, Bookman AA, Zev Shainhouse J. Efficacy and safety of topical diclofenac containing dimethyl sulfoxide (DMSO) compared with those of topical placebo, DMSO vehicle and oral diclofenac for knee osteoarthritis. Pain. 2009; 143(3):238-245

90.   Strand V, Bergman M, Singh JA, Gibofsky A, Kivitz A, Young C. Low-dose SoluMatrix diclofenac in patients with osteoarthritis pain: impact on quality of life in a controlled trial. Clinical Rheumatology. 2017; 36(6):1357-1367

91.   Stravitz RT, Kramer DJ. Chapter 20 - Acute Liver Failure. 'In:' Boyer TD, Manns MP, Sanyal AJ, editors. Zakim and Boyer's Hepatology (Sixth Edition). Saint Louis: W.B. Saunders. 2012. p. 327-351.

92.   Truitt KE, Sperling RS, Ettinger WH, Jr., Greenwald M, DeTora L, Zeng Q et al. A multicenter, randomized, controlled trial to evaluate the safety profile, tolerability, and efficacy of rofecoxib in advanced elderly patients with osteoarthritis. Aging-Clinical & Experimental Research. 2001; 13(2):112-121

93.   Villar E, Remontet L, Labeeuw M, Ecochard R. Effect of age, gender, and diabetes on excess death in end-stage renal failure. Journal of the American Society of Nephrology. 2007; 18(7):2125-2134

94.   Wadsworth LT, Kent JD, Holt RJ. Efficacy and safety of diclofenac sodium 2% topical solution for osteoarthritis of the knee: a randomized, double-blind, vehicle-controlled, 4 week study. Current Medical Research and Opinion. 2016; 32(2):241-250

95.   Wailoo A, Hernandez Alava M, Escobar Martinez A. Modelling the relationship between the WOMAC Osteoarthritis Index and EQ-5D. Health Qual Life Outcomes. 2014; 12:37

96.   Ward A, Bozkaya D, Fleischmann J, Dubois D, Sabatowski R, Caro JJ. Modeling the economic and health consequences of managing chronic osteoarthritis pain with opioids in Germany: comparison of extended-release oxycodone and OROS hydromorphone. Current Medical Research and Opinion. 2007; 23(10):2333-2345

97.   Ward S, Lloyd Jones M, Pandor A, Holmes M, Ara R, Ryan A et al. A systematic review and economic evaluation of statins for the prevention of coronary events. Health Technology Assessment. 2007; 11(14):1-160, iii-iv

98.   Watson J, Helliwell P, Morton V, Adebajo A, Dickson J, Russell I et al. Shoulder acute pain in primary healthcare: is retraining effective for GP principals? SAPPHIRE--a randomized controlled trial. Rheumatology. 2008; 47(12):1795-1802

99.   Weiler N, Schlotmann A, Schnitzbauer AA, Zeuzem S, Welker MW. The Epidemiology of Acute Liver Failure. Dtsch Arztebl Int. 2020; 117(4):43-50

100.   Zhao SZ, Dedhiya SD, Bocanegra TS, Fort JG, Kuss ME, Rush SM. Health-related quality-of-life effects of oxaprozin and nabumetone in patients with osteoarthritis of the knee. Clinical Therapeutics. 1999; 21(1):205-217

Appendices

Appendix A: Probabilistic analysis input parameters

The tables below summarise all probabilistic inputs in the model and the distribution parameters used.

Table 45: Probabilities, rate and utilities

Parameter	Mean	Standard error	Distribution	Parameters of the beta distribution (except where stated)	
				alpha	beta
Baseline events					
Constipation	0.064		Beta	75	1170
Nausea	0.074		Beta	86	1170
Vomiting	0.032		Beta	37	1170
Dyspepsia	0.022		Beta	17	773
Symptomatic ulcer	0.004		Beta	3	773
gastrointestinal bleeds	0.001	0.001	Beta	0.52	645
AKI	0.002147		Beta	997853	0.002
Progression to CKD stage 3-4	0.194215		Beta	47	195
Progression to CKD stage 5 <69 years	0.002		Beta	5	3041
Progression to CKD stage 5 70-79 years	0.001		Beta	3	3043
Progression to CKD stage 5 >79 years	0.001		Beta	2.32	3044
Vertigo	0.078		Beta	91	1170
Drug-related adverse events					
Paracetamol					
Hepatic events	5.97	0.49[a]	Lognormal		
Oral NSAIDs plus PPI					
Gastrointestinal (bleeding or perforation)	2.32	0.62[a]	Lognormal		
Cardiovascular	2.51	0.63[a]	Lognormal		
Renal	2.67	0.95[a]	Lognormal		
Oral NSAIDs alone					
Gastrointestinal (bleeding or perforation)	4.57	0.10[a]	Lognormal		

Parameter	Mean	Standard error	Distribution	Parameters of the beta distribution (except where stated)	
				alpha	beta
Gastrointestinal (non-bleeding)	1.17	0.56[a]	Lognormal		
Cardiovascular	1.15	0.14[a]	Lognormal		
Renal	1.96	0.67[a]	Lognormal		
Strong opioids					
Gastrointestinal (non-bleeding)	1.63	0.36[a]	Lognormal		
Central nervous system	1.93	0.07[a]	Lognormal		
Buprenorphine					
Gastrointestinal (non-bleeding)	2.26	0.19[a]	Lognormal		
Central nervous system	2.48	0.24[a]	Lognormal		
Mortality					
Hip fractures (men)	1.9	0.096	Lognormal		
Hip fractures (women)	1.86	0.077	Lognormal		
Utility multipliers					
Nausea	0.733	0.055	Beta	47	17
Vomiting	0.733	0.055	Beta	47	17
Dyspepsia	0.733	0.055	Beta	47	17
Symptomatic ulcer	0.552	0.048	Beta	60	49
gastrointestinal bleeds	0.459	0.049	Beta	46	55
Stable Angina (New Event)	0.808	0.038	Beta	86	20
Stable Angina (Post Event)	0.808	0.038	Beta	86	20
Unstable Angina (New Event)	0.770	0.038	Beta	94	28
Unstable Angina (Post Event)	0.808	0.038	Beta	86	20
MI (New Event)	0.760	0.018	Beta	427	135
MI (Post Event)	0.880	0.018	Beta	286	39
TIA (New Event)	0.900	0.025	Beta	129	14
TIA (Post Event)	0.900	0.025	Beta	129	14
Stroke (New Event)	0.628	0.040	Beta	91	54
Stroke (Post Event)	0.628	0.040	Beta	91	54

Parameter	Mean	Standard error	Distribution	Parameters of the beta distribution (except where stated)	
				alpha	beta
Heart Failure (New event)	0.683	0.020	Beta	369	171
Heart Failure (Post Event)	0.683	0.020	Beta	369	171
Vertigo	0.733	0.055	Beta	47	17
AKI	0.525	0.033	Beta	121	110
Moderate CKD	0.861	0.02[a]	Beta		
Severe CKD	0.798	0.03[a]	Beta		
ALF	0.525	0.033	Beta		
Liver transplant	0.800	0.011	Beta	1077	269
Liver transplant (post event)	0.850	0.004	Beta	7124	1257
Utility decrements					
Constipation	0.072	0.02	Gamma	16	0.0045
Hip fracture	0.237	0.074	Gamma	10	0.0231
Post-Hip fracture 12 ms	0.220	0.074	Gamma	8.8	0.0249
Hip - post hip	0.017	0.10	Gamma	0.03	0.6470
Falls emergency	0.014	0.005	Gamma	8.0	0.0017
Falls hospitalisation	0.144	0.210	Gamma	0.5	0.3054
Falls admitted residential care	0.060	0.253	Gamma	0.1	1
Costs (annualised)					
Stable Angina (New Event)	776	155	Gamma	25	31
Stable Angina (Post Event)	359	67	Gamma	29	12
Unstable Angina (New Event)	2,385	53	Gamma	2025	1
Unstable Angina (Post Event)	359	67	Gamma	29	12
MI (New Event)	4,679	112	Gamma	1745	2
MI (Post Event)	1,010	170	Gamma	35	28
TIA (New Event)	1,682	76	Gamma	490	3
TIA (Post Event)	144	161	Gamma	1	180
Stroke (New Event)	18,126	3,625	Gamma	25	725
Stroke (Post Event)	6,914	1,383	Gamma	25	276

Parameter	Mean	Standard error	Distribution	Parameters of the beta distribution (except where stated)	
				alpha	beta
Heart failure (New Event)	2,674	104	Gamma	661	4
Heart Failure (Post Event)	928	271	Gamma	12	79
Hip fracture (acute)	14,971	268	Gamma	3118	4
Post-hip fracture	2,260	95	Gamma	571	3

(a) Log scale