NCC-WCH

Version 1.4

Menopause

Appendix L to M

Clinical guideline

Methods, evidence and recommendations

25 September 2015

Final

Commissioned by the National Institute for Health and Care Excellence

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Funding
Registered charity no. 213280

Contents

Appendices

Appendix A: Scope

The scope is presented in a separate document

Appendix B: Stakeholders

The list of stakeholders is presented in a separate document

Appendix C: Declarations of interest

The declarations of interests are presented in a separate document

Appendix D: Review protocols

The review protocols are presented in a separate document

Appendix E: Search strategies

The search strategies are presented in a separate document

Appendix F: PRISMA flow charts

The PRISMA flow charts are presented in a separate document

Appendix G: Excluded studies

The excluded studies are presented in a separate document

Appendix H: Evidence tables

The evidence tables are presented in a separate document

Appendix I: GRADE profiles

The GRADE profiles are presented in a separate document

Appendix J: Forest plots

The forest plots are presented in a separate document

Appendix K: Network meta-analysis of interventions in the pharmacological and non-pharmacological treatment of short term symptoms for women in menopause

The network-meta analysis is presented in a separate document.

Appendix L: Health economics

L.1 What is the most clinical and cost-effective treatment for the relief of individual menopause-related symptoms for women at menopause?

L.1.1 Introduction

This section provides details of the review of published health economic literature and the health economic modelling undertaken for this guideline.

L.1.2 Review of the literature

A search of economic evidence relating to short term treatments for menopause symptoms identified 480 papers. After screening titles and abstracts 41 full text articles were retrieved for further review. Of these 41 studies, 9 were considered to be relevant to the review question and are summarised in Table 1.32 studies were excluded, the majority of which did not provide an economic evaluation of alternative treatment options. Other reasons for exclusion were that the studies did not include relevant interventions, were on the wrong population or focused on a longer timeframe.

The economic evaluations considered various hormone replacement therapies (combined oestrogen and progesterone or oestrogen alone) and tibolone. All studies employed a cost-utility analysis and used a Markov model except one (Ylikangas 2005), which was a trial-based economic evaluation.

- Three studies were carried out in the US (Botteman 2004; Col 2004; Lekander 2005).
- Two studies were carried out in the UK (Lekander 2009; Swift 2005)
- One Swedish studies (Zethraeus 2005)
- Three Canadian studies (Coyle 2003; Brown 2006; Diaby 2007)
- One Finnish trial-based economic evaluation (Ylikangas 2005)

Economic evaluations that included other, non-hormonal, preparations were not identified.

A cost-utility analysis (Botteman 2004) compared two preparations of continuous combined HRT (1mg of norethindrone acetate/5 µg of ethinyl estradiol (NA/EE) and 0.625 mg/day of conjugated estrogens plus 2.5 mg of medroxyprogesterone (CEE/MPA) versus no therapy for the management of vasomotor symptoms, and including the impact on breakthrough bleeding. Women between 45-55 years old with an intact uterus and moderate-severe vasomotor symptoms were included in this study and data were taken from a clinical trial (Speroff 2000). The Markov model had a 1 year time horizon and 3 month time cycle. For the model, 34% elimination in hot flushes every 3 months was assumed. Utility data for vasomotor symptoms were taken from two published studies (Daly 1993; Zethraeus 1997) and averaged for ‘moderate-severe’ symptoms. Data for the ‘well’ health state were derived from a study using the quality of wellbeing scale. In a sensitivity analysis the time horizon was extended to 3 years and the utility for ‘mild’ state was assigned. A probabilistic sensitivity analysis was also carried out. The results show that NA/EE was the most cost-effective intervention dominating CEE/MPA and with an incremental cost-effectiveness ratio (ICER) 6,200 USD relative to no therapy and the authors reported that this finding was supported by the sensitivity analysis.

The cost-effectiveness of HRT therapy versus placebo was assessed for an average population of Swedish women with menopausal symptoms (Zethraeus 2005). The analysis included data from the Women's Health Initiative study. The cost-effectiveness of HRT was calculated in six patient groups dependent on age commencing treatment (50, 55 or 60 years old) and uterine status (intact uterus or hysterectomised). The Markov model had a 50-year time horizon divided into cycle lengths of 1 year. The model included disease states related to coronary heart disease, stroke, venous thromboembolic events, breast cancer, colorectal cancer, hip fracture, vertebral fracture, and wrist fracture. Quality of life scores were taken from the literature (Zethraeus 2002), with an average gain of 0.29 from HRT treatment calculated in a Swedish time-trade off preference elicitation (Zethraeus 1997). Compared to no treatment, this study found that HRT was a cost-effective strategy with an ICER of 12,807 SEK per QALY in women with an intact uterus and with an ICER of 8,266 SEK per QALY in women who had received a hysterectomy. A number of one-way sensitivity analyses were run and the authors reported that HRT would remain cost-effectiveness providing quality of life scores exceeded 0.013.

A Canadian health technology assessment (Brown 2006) conducted an economic analysis comparing the cost-effectiveness of transdermal HRT against oral HRT and against placebo for women with post-menopausal symptoms. Using a Markov model, the cost-effectiveness was assessed through the impact on menopausal symptoms and vaginal bleeding over five years, with a Markov cycle length of 3 months. The model consisted of four main outcomes to generate QALYs: compliance, bleeding, postmenopausal symptoms and mortality. Postmenopausal symptoms were measured through symptom severity, with mild symptoms classed as less than 3 hot flushes per day and severe symptoms classed as 10 or more a day. Utilities for postmenopausal symptoms were taken from the time-trade off study (Daly 1993), with vaginal bleeding utilities generated from a standard gamble of a subsample of osteoporosis patients in a Canadian hospital (Cranney 2001). A two-year treatment timeframe was assessed in a sensitivity analysis. The authors reported that transdermal HRT patches were not cost-effective relative to oral HRT for either the moderate or severe post-menopausal symptom groups. Relative to no treatment, transdermal patches had an incremental cost per QALY of approximately 32,300 CAD for the patients with moderate symptoms. For women with severe symptoms, relative to no treatment, the cost per QALY gained was approximately 8,300 CAD.

A trial based economic evaluation was undertaken in postmenopausal women in Finland treated for up to nine consecutive years (Ylikangas 2007). The evaluation compared two continuous combined therapies for women with menopausal symptoms (1mg Oestradiol valerate and 2.5mg medroxyprogesterone; 2mg Oestradiol valerate and 5mg medroxyprogesterone) and applied data from the general Finnish population as a control group. Key events associated with use of HRT included were myocardial infraction, stroke, deep vein thrombosis, pulmonary embolism, colon cancer, breast cancer, cervical and other gynaecological cancers, and fractures of the vertebrae, radius and hip. Health related quality of life was collected using the 15D scale during the trial, a Finnish preference-based health utility instrument. The authors conclude that continuous combined HRT is cost-effective for up to 9 years, with an ICER of 4613 Euros per QALY, for non-hysterectomised women predominantly in the age range of 55-64 years who are experiencing climacteric symptoms. The results of the analysis were not sensitive to a number of one-way sensitivity analyses, but the authors noted concerns about the generalizability of the trial participants to the Finnish general population.

An economic evaluation undertaken from a UK NHS perspective compared combined 1 mg estradiol and 0.5 mg norethisterone versus no therapy for the treatment of menopausal symptoms in women with an intact uterus and estradiol alone for hysterectomised women (Lekander 2009a). Relative risks of various outcomes were based on the Women's Health Initiative (WHI) study. The Markov model had a lifetime horizon and a yearly time cycle, with treatment duration set to 5 years. As well as improvement in menopausal symptoms the

study also assessed the effect of HRT on risk of stroke, venous thromboembolism, breast cancer, colorectal cancer, hip fracture, vertebral fracture, wrist fracture and coronary heart disease. Quality of life data for the disease risks were based on EQ-5D and taken from other studies. Gain in quality of life from menopausal relief with HRT was based on published time trade off values (Zethraeus 1997). The authors reported that treatment with HRT for menopausal symptoms was cost-effective in both groups of women. The same authors used a similar approach to compare HRT against no therapy in a US setting (Lekander 2009b). Again therapy was compared in two population groups, women with an intact uterus and hysterectomised and the authors reported that HRT was cost-effective in women with menopausal symptoms.

A Canadian study compared continuous combined therapy 1mg norethindrone and 5mcg ethinyloestradiol versus 0.625mg conjugated equine oestrogen and 2.5mg medroxyprogesterone acetate versus no therapy (Coyle 2003). The Markov model incorporated the presence or absence of vaginal bleeding, menopausal symptoms and hip fracture in a 50 year old menopausal woman. A lifetime model with a 3 month Markov time cycle was employed with a third party payer perspective. Treatment was assumed for 5 years but a sensitivity analysis using a 1 year course of treatment was undertaken. Estimates of treatment efficacy were based on reduction in frequency of hot flushes and night sweats. Utility values for menopausal symptoms were based on previously published values (Daly 1993). The authors concluded that 1mg norethindrone and 5mcg ethinyloestradiol was the most cost-effective intervention with the authors additionally reporting that sensitivity analysis confirmed the robustness of the initial results.

A Markov model was used to compare a 3 year treatment course of synthetic hormone tibolone 2.5mg versus conjugated equine oestrogens 0.625mg with medroxyprogesterone acetate 2.5mg (CEE/MA) in postmenopausal women (Diaby 2007). The model structure was based on a previously published study (Coyle 2003) with Markov cycles of 3 months duration. The analysis incorporated persistence with treatment, vaginal bleeding and climacteric symptoms (including hot flushes, night sweats, mood changes, sexual dysfunction). The study, set in Canada, used a third party payer perspective (Quebec healthcare system). Utility data were based on previously used values (Botteman 2004). Markov transition probabilities were based on a published study (Hammar 1998). The authors concluded that tibolone is a cost-effective alternative to CEE/MA. Univariate and bivariate sensitivity analysis was undertaken with the authors concluding that different input values would not significantly change the conclusions of the study.

In a UK study, the cost-effectiveness of low dose 0.3mg conjugated oestrogen and 1.5mg medroxyprogesterone acetate injection (0.3/ 1.5mg CE/MPA) versus a higher dose 0.625mg conjugated oestrogen and 5mg medroxyprogesterone acetate injection (0.625/ 5mg CE/MPA) was compared in postmenopausal women with an intact uterus (Swift 2005). The evaluation was done from an NHS perspective and utilised a Markov model approach incorporating bleeding, breast pain, breast symptoms, vaginal candidiasis and treatment discontinuation rates. Treatment was evaluated for a period of one-year and health state utilities for mild, moderate and severe menopausal symptoms were derived from previously published work (Daly 1993). Sensitivity analyses were carried out, including a probabilistic sensitivity analysis. The findings show that compared to the high dose treatment, the low dose 0.3/ 1.5mg CE/MPA is the most cost-effective treatment.

Table 1: Profile of health economic studies

Study	Limitations	Applicability	Other comments	Incremental			Uncertainty
				Costs	Effects	ICER	
Botteman 2004	Transition probabilities for vasomotor symptoms derived from a trial with a small sample size Did not account for long-term clinical or economic aspects	Partially applicable (US study)	Study used a Markov decision-analytic model with a 1-year time horizon Research sponsored in part by Pfizer	NA/EE vs no therapy $680.84 CEE/MPA vs no therapy $847.93	NA/EE vs no therapy 0.110 QALYs CEE/MPA vs no 0.104 QALYs	NA/EE dominates CEE/MPA NA/EE vs no therapy $6,200 per QALY CEE/MPA v no therapy $8,200 per QALY	Univariate, bivariate, threshold and probabilistic sensitivity analysis
Brown 2006	Hot flushes used as proxy for presence and severity of postmenopausal symptoms	Partially applicable (Canadian study)	Study employed a Markov decision-analytic model with a 5-year time horizon	Patch vs oral $296 Patch vs no therapy $654-665	Patch vs oral 0.00 QALYs Patch vs no therapy 0.02-0.08 QALYs	Oral dominates patch Patch compared to no therapy for moderate ($32,300 per QALY) and severe ($8,300 per QALY)	One-way and probabilistic sensitivity analysis undertaken
Coyle 2003	Hot flushes used as proxy for menopausal symptoms No probabilistic sensitivity analysis conducted	Partially applicable (Canadian study)	Study employed a Markov decision-analytic model with a 5-year time horizon Study funded by Pfizer inc.	1st line: CEE/MPA vs no therapy 100 CAD NA/EE vs CEE/MPA 600 CAD	1st line: CEE/MPA vs no therapy 0.30 QALYs NA/EE vs CEE/MPA 0.03 QALYs	1st line: CEE/MPA vs no therapy 333 CAD per QALY NA/EE vs CEE/MPA 20,300 CAD per	One-way and threshold sensitivity analysis undertaken

Study	Limitations	Applicability	Other comments	Incremental			Uncertainty
				Costs	Effects	ICER	
						QALY 2nd line: CEE/MPA vs no therapy CEE/MPA dominates NA/EE vs CEE/MPA 16,400 CAD per QALY	
Lekander 2009[a]	No comparison with alternative treatment No probabilistic sensitivity analysis conducted	Directly applicable (UK study)	Study employed a Markov decision analytic model with a lifetime horizon Study funded and conducted by consultants for Wyeth	Women with an intact uterus: HRT vs No therapy 677 GBP Hysterectomised women: HRT vs No therapy 252 GBP	Women with an intact uterus: HRT vs No therapy 1.17 QALYs Hysterectomised women: HRT vs No therapy 1.23 QALYs	Women with an intact uterus: HRT v no therapy 580 GBP per QALY Hysterectomised women: HRT vs No therapy 205 GBP per QALY	Univariate and threshold sensitivity analysis undertaken
Lekander 2009[b]	No comparison with alternative treatment No probabilistic sensitivity analysis conducted	Partially applicable (US study)	Study employed a Markov decision analytic model with a lifetime horizon Study funded and conducted by consultants	Women with an intact uterus: HRT vs No therapy 3224 USD	Women with an intact uterus: HRT vs No therapy 3224 USD	Women with an intact uterus: HRT v no therapy 2803 USD per QALY	Univariate and threshold sensitivity analysis undertaken

Study	Limitations	Applicability	Other comments	Incremental			Uncertainty
				Costs	Effects	ICER	
	Study conducted from a societal perspective		for Wyeth	Hysterectomised women: HRT vs No therapy 358 USD	Hysterectomised women: HRT vs No therapy 358 USD	Hysterectomised women: HRT vs No therapy 295 USD per QALY	
Swift 2005	Model structure and type presented unclearly. Utilities on menopausal symptom severity only included	Directly applicable (UK study)	Study developed an economic model over a one-year time horizon Study funded and conducted by consultants for Wyeth	Low-dose vs high dose CE/MPA -£1,443	Low-dose vs high dose CE/MPA 0.62-1.49 QALYs	Low dose dominates high dose CE/MPA	Probabilistic sensitivity analysis undertaken
Yilkangas 2007	No probabilistic sensitivity analysis conducted	Partially applicable (Finnish study	Study conducted a trial-based economic evaluation over a 9-year time horizon Study was funded by Orion Pharma	ccHRT vs gen population €101	ccHRT vs gen population 0.022 QALYs	ccHRT vs gen population €4613 per QALY	Univariate sensitivity analysis undertaken
Zethraeus 2005	Study conducted from a societal perspective No probabilistic sensitivity analysis undertaken	Partially applicable (Swedish study)	Study employed a Markov decision analytic model with a lifetime horizon Funding for this study was provided by Wyeth Lederle	Intact uterus HRT vs No HRT SEK 15,242 Hysterectomised HRT vs No HRT SEK 10,107	Intact uterus HRT vs No HRT 1.19 QALYs Hysterectomised HRT vs No HRT 1.22 QALYs	Intact uterus HRT vs No HRT SEK 12,807 per QALY Hysterectomised HRT vs No HRT SEK 8,266 per QALY	Univariate sensitivity analysis undertaken

Study	Limitations	Applicability	Other comments	Incremental			Uncertainty
				Costs	Effects	ICER	
Diaby 2007	Assumptions made concerning utility of reduction of symptoms No probabilistic sensitivity analysis	Partially applicable (Canadian study)	Study employed a Markov decision-analytic model with a 3-year time horizon	Tibolone (2.5mg) vs ccHRT (CEE/MPA 0.625/2.5mg) $253	Tibolone (2.5mg) vs ccHRT (CEE/MPA 0.625/2.5mg) 0.03 QALYs	Tibolone (2.5mg) vs ccHRT (CEE/MPA 0.625/2.5mg) $9,198 per QALY	Univariate and bivariate sensitivity analysis undertaken

L.1.3 Methods

No published health economic literature was identified that addressed the breadth of treatment alternatives included in the network meta-analysis for this guideline and it was therefore considered important to develop a de Novo model which reflected this approach to synthesising clinical effectiveness data.

A semi-Markov decision analytic model was developed in Microsoft Excel® to assess the cost effectiveness of 5 years of use of HRT, non-HRT drugs, herbal preparations, and other interventions given to menopausal women with vasomotor symptoms starting treatment at 50 years of age, reflecting the average age at which women typically start the menopause.

The model was run for three populations:

1. Women with a uterus
2. Women without a uterus
3. Women who have had breast cancer or are at high risk of breast cancer.

To reflect uncertainty in model parameters, the results were assessed using probabilistic sensitivity analysis. The model aimed to follow the NICE reference case unless otherwise stated.

L.1.3.1 Time horizon

Short-term hormonal replacement therapy has been defined as the use of oestrogen or an oestrogen-progesterone combination to treat menopause symptoms for the shortest possible time and with the lowest possible dose (Department of Health, Social Services and Public Safety, 2003) consistent with treatment goals. This would usually be for a period 2-5 years and therefore this model used to inform guideline recommendations is based on a woman taking treatment for menopausal symptoms for five years, the maximum period that would normally be considered short-term.

L.1.3.2 Clinical outcomes included in the model

Treatment is intended to reduce vasomotor symptoms (hot flushes), but can cause vaginal bleeding as a side-effect, and this will impact the discontinuation rate. As part of the protocol for the clinical review the GDG prioritised the outcomes listed in **Table 2** below.

Table 2: GDG prioritised outcomes

Outcome	Included in network meta-analysis
Vasomotor	
Frequency of hot flushes (including night sweats)	Yes
Altered sexual function	
Frequency of sexual intercourse	No
Psychological symptoms	
Anxiety	No
Depression	No
Musculoskeletal symptoms	
Symptom relief	No
Muscle strength	No
Safety outcomes	
Discontinuation	Yes

Outcome	Included in network meta-analysis
Major adverse events	
Breast cancer	No
Other cancer	No
Arterial disease	No
Venous thromboembolic disease	No
Fracture	No
Mortality	No
Minor adverse events	
Vaginal Bleeding pattern	Yes

After discussion with the GDG it was agreed that the following outcomes would be incorporated within the health economic model:

- Vasomotor symptoms
- Vaginal bleeding (not included as an outcome for women without a uterus)
- Discontinuation of treatment
- Breast cancer (not included as an outcome for women with breast cancer)
- Venous thromboembolism (VTE)

The first three outcomes above reflect the network meta-analyses that was undertaken for this guideline (see Appendix K). The output from the network meta-analysis can be considered as representing the “gold standard” measure of treatment effectiveness.

Vasomotor symptoms, or hot flushes, are the most commonly reported menopausal symptom and are therefore frequently reported as an outcome in intervention studies. The model used the frequency of vasomotor symptoms measured as mean number of flushes per day. Severity of symptoms was not considered as part of this outcome due to the variation in scores used to measure them. Vaginal bleeding and discontinuation were both included as adverse events of treatment.

Whilst the model considered only short term use of hormone replacement therapy for menopausal symptoms it is recognised that there are potentially longer term consequences of short term use of hormone replacement therapy and it is important that such factors are not overlooked in an assessment of cost-effectiveness.

The GDG’s view after reviewing the evidence was that both breast cancer and VTE were important long term adverse effects of short term treatment and should be included in the model. More specifically, the evidence from the clinical review undertaken for this guideline of the longer term effects of combined HRT administered for menopausal symptoms on incidence of breast cancer found that there was an increased risk of breast cancer with duration of use, but that this risk reduced after stopping HRT. The clinical review of the effect of combined HRT or oestrogen alone on the risk of VTE found that the risk of developing VTE increases with age and in the presence of other factors such as obesity, smoking or the presence of an inherited thrombophilia. The increase in risk of VTE occurs rapidly after starting HRT in oral (tablet) formulations and continues until discontinuation. Oral (tablet) oestrogen alone or combined HRT increases the risk of VTE and this occurs immediately after starting treatment although no such increase was observed in women using normal dose transdermal preparations.

The review undertaken for this guideline of CHD related to HRT use found no convincing evidence that administration of HRT increases risk in women under 65 years of age. This was evident for oestrogen and oestrogen plus progestogen preparations. There was evidence that HRT increases the risk of stroke when administered orally, however the absolute risk was very small. The clinical review of the effect of HRT on subsequent

development of osteoporosis found HRT use was beneficial for the duration of therapy, but decreased following cessation. Furthermore, the risk of fractures in women around the age of menopause is very low. Consequently, the GDG agreed that the omission of these outcomes from the model was reasonable.

L.1.3.3 Interventions and comparisons

The treatment alternatives for each of the three different populations were generally determined by the treatments that were included in the network meta-analyses for these populations as listed in Table 3, Table 4 and Table 5. This was decided because the network meta-analysis provided only results on the comparative effectiveness of the interventions that were included in the network.

A class analysis was adopted for the network meta-analysis where different dosages of pharmacological treatment (categorized as low, medium and high) were grouped under the same class. However, oral and transdermal preparations of oestrogen alone and oestrogen plus progesterone were fitted in the network as a separate class. The health economic analysis assessed interventions according to how they were grouped for the purposes of the network meta-analysis.

A no treatment alternative, based on baseline data, was included in all analyses.

Women with a uterus

Table 3: Classes of interventions included in the network meta-analyses for women with a uterus

Intervention	Vasomotor symptoms[a]	Bleeding[a]	Discontinuation[a]
Acupuncture	✓	No NMA[d]	No NMA[e]
Bazedoxifene with conjugated oestrogen	No NMA	No NMA	✓
Chinese herbal medicine	✓	No NMA[d]	✓
Gabapentin	No NMA[c]	✓	✓
Isoflavones/Genistein/Soy	✓	No NMA[d]	✓
Multibotanicals	✓	No NMA[d]	✓
Oestradiol + Progestogen non-oral	✓	No NMA[b]	No NMA[b]
Oestradiol + Progestogen oral	✓	✓	✓
Raloxifene	✓	No NMA	No NMA
Black cohosh	✓	No NMA[d]	✓
Valerian root	No NMA[d]	No NMA[d]	✓
Sham acupuncture	✓	No NMA	No NMA
SSRIs/SNRIs	✓	✓	✓
Tibolone	✓	✓	✓

(a) Data available for outcome for this intervention unless otherwise stated
(b) No NMA data – assumed the same as oral oestradiol + Progestogen in women with a uterus
(c) No NMA data – assumed to be the same as Gabapentin in women with breast cancer
(d) No NMA data – assumed to be the same as baseline
(e) No NMA data – assumed not applicable

The GDG decided that sham acupuncture and raloxifene, whilst providing indirect evidence to the network, should not be included in the health economic model. Sham acupuncture, often used as a placebo to test acupuncture efficacy, was not considered a viable treatment option to be offered on the NHS and raloxifene, a selective oestrogen receptor modulator (SERM), although known to reduce the risk of breast cancer, it is considered to worsen vasomotor symptoms. Bazedoxifene with conjugated oestrogen was also not included in the

analysis as there was only network meta-analysis data for discontinuation. The full list of comparators included in the analysis for women with a uterus was as follows:

- No treatment
- Acupuncture
- Chinese herbal medicine
- Gabapentin
- Isoflavones/Genisten/Soy
- Multibotanicals
- Oestradiol + progestogen non-oral
- Oestradiol + progestogen oral
- Black cohosh
- Valerian root
- SSRIs/SNRIs
- Tibolone

Women without a uterus

Table 4: Classes of interventions included in the network meta-analyses for women without a uterus

Intervention	Vasomotor symptoms[a]	Discontinuation[a]
Acupuncture	✓	No NMA[e]
Chinese herbal medicine	✓	✓
Gabapentin	No NMA[b]	✓
Isoflavones/Genistein/Soy	✓	✓
Multibotanicals	✓	✓
Oestradiol alone non-oral	No NMA[c]	✓
Raloxifene	✓	No NMA
Black cohosh	✓	No NMA
Valerian root	No NMA[d]	✓
Sham acupuncture	✓	No NMA
SSRIs/SNRIs	✓	✓

(a) Data available for outcome for this intervention unless otherwise stated
(b) No NMA data - assumed to be the as Gabapentin in women with breast cancer
(c) No NMA data – assumed to be the same as oestradiol and progestogen non-oral in women with a uterus
(d) No NMA data – assumed to be the same as baseline
(e) No NMA data – assumed not applicable

Similarly to the population of women with a uterus, the GDG decided that sham acupuncture and raloxifene were not appropriate comparators to include in the health economic model for women without a uterus. In addition the GDG requested that oral oestradiol should be included as it is the most common type of HRT for women without uterus. As this treatment was not connected in the network and therefore not considered in this model of network meta-analysis, it was assumed that oral oestradiol alone in women without a uterus would have the same relative treatment effects as oral oestradiol plus progestogen in women with a uterus.

Therefore the complete list of interventions compared in the analysis for this population was as follows:

- No treatment
- Acupuncture

- Chinese herbal medicine
- Gabapentin
- Isoflavones/Genistein/Soy
- Multibotanicals
- Oestradiol alone
- Oestradiol alone non-oral
- Black cohosh
- Valerian root
- SSRIs/SNRIs

Women with breast cancer

Table 5: Classes of interventions included in the network meta-analyses for women without a uterus

Intervention	Vasomotor symptoms[a]	Bleeding[a]	Discontinuation[a]
Gabapentin	✓	No NMA[b]	✓
Isoflavones/Genistein/Soy	✓	No NMA[c]	✓
SSRIs/SNRIs	✓	No NMA[c]	✓
St John's Wort	✓	No NMA[c]	No NMA[c]

(a) Data available for outcome for this intervention unless otherwise stated
(b) No NMA data - assumed to be the as Gabapentin in women with uterus
(c) No NMA data – assumed to be the same as baseline

All interventions included in the network meta-analysis were included in the health economic analysis for this population along with no treatment.

L.1.3.4 Basic model structure

The model structure can be categorised as being of a semi Markov decision analytic type. The decisions represent the alternative interventions, including no treatment, that are being compared. For each decision the costs and benefits, measured in QALYs, are calculated based on the probabilities of various events and outcomes reflecting the comparative risks and benefits of the treatment alternatives derived from the evidence.

A Markov model involves the transition of a hypothetical patient across different 'health states' over time, divided into equally spaced cycles. The health states in this model are "continue on treatment", "discontinue treatment" and "death", which is known as an "absorbing state" as there can be no transition to an alternative state once this state is entered. Within each state costs and utilities are assigned according to the probabilities associated with the health state decision sub-tree (the various events and outcomes that occur within cycle)

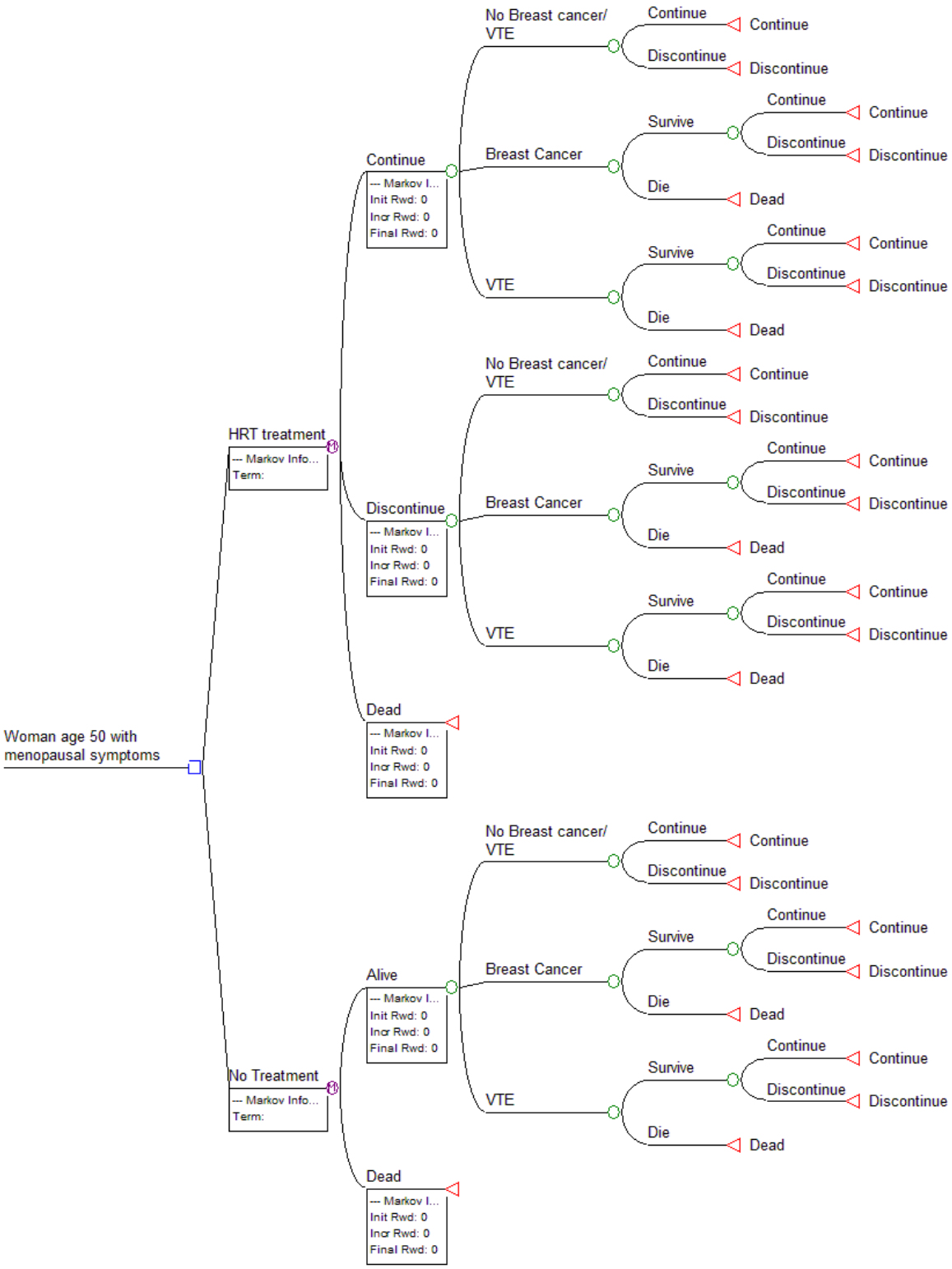
Each cycle in this model represents three months and therefore to represent the 5-year timeframe there are 20 cycles in total. Transition between different states occurs at the end of cycles and is determined by transition probabilities derived from the literature, the network meta-analysis or assumption. A schematic of the Markov model is shown in Figure 1:
Markov decision tree for short-term treatment of menopause symptoms

. For ease of illustration only two alternatives are depicted and the "HRT Treatment" branch gives the tree structure for all active treatment alternatives.

Breast cancer and VTE probabilities were estimated using baseline data (see Appendix L.1.3.8), a meta-analysis of relative risks that was undertaken as part of the clinical review for this guideline and from the literature. Discontinuation during the first two Markov cycles (6

months) is based on the network meta-analysis and a range of assumptions thereafter. For HRT treatments the network meta-analysis included studies that had a measure of discontinuation at between 12 weeks and 6 months. For non-HRT treatments the network meta-analysis included studies that had a measure of discontinuation at between 4 weeks and 6 months. For included studies that reported only for a time period of less than 6 months, there is an implicit assumption within the network meta-analysis that no further discontinuation would take place between the last measured time point and six months. This assumption is supported to some extent by studies which reported discontinuation at multiple time points. The model additionally makes a simplifying assumption that the discontinuation rate is the same across both cycles, which still gives a higher absolute rate of discontinuation in the first cycle.

Figure 1: Markov decision tree for short-term treatment of menopause symptoms



Vasomotor symptoms and bleeding were modelled separately from the Markov process using data from the network meta-analysis. Vasomotor symptoms and bleeding are assumed to be constant for the 5 year duration of the model although the duration of bleeding can be amended as part of a sensitivity analysis (see also L.1.3.6).

L.1.3.5 Probabilistic sensitivity analysis

In reporting clinical effectiveness it is usual and good practice to take into account the uncertainty of a relative treatment effect by reporting confidence intervals around the point estimate. Similarly, in health economic analysis it is important to take into account the uncertainty around model inputs. This can sometimes be achieved through one way sensitivity analysis, where one input value is altered at a time in order to assess what change that input has on the model's results. However, whilst that can often provide useful insights into what inputs are driving the models results it is inadequate to address the uncertainty which exists simultaneously across all model inputs.

Probabilistic sensitivity analysis, using Monte Carlo simulation techniques, allows for uncertainty across all model inputs to be addressed. Simulation involves running the model many times. In each simulation, rather than using the point estimate of the input, the value is sampled from its probability distribution. For inputs that are based on a large sample the probability distribution will be relatively narrow and the sampled inputs will reflect that. This model assessed the cost-effectiveness of the various treatment alternatives using probabilistic sensitivity analysis.

In undertaking the network meta-analysis a large number of simulations or iterations of relative treatment effect are performed accounting for the within-study correlation between treatment effects induced by multi-arm trials. Up to a 100,000 samples of the relative treatment effect derived from the network meta-analysis are used in the probabilistic sensitivity analysis for the outcomes of discontinuation, bleeding (when relevant) and vasomotor symptoms.

Baseline risks of clinical outcomes, relative risks for breast cancer and venous thromboembolism and some costs were also sampled probabilistically. The model included some deterministic inputs, such as costs based on published prices for example. Health state utilities were also deterministic inputs in the model as, given the way they were estimated, it was difficult to define a meaningful distribution from which to sample. However, to address this limitation in the model, extensive one way sensitivity analysis was undertaken on those variables influencing QALY gain to assess the extent to which cost-effectiveness was influenced by changes to these inputs.

L.1.3.6 Outcome modelling assumptions

A number of assumptions and simplifications were made in modelling the different clinical outcomes included in this model. These assumptions and their rationale is described below. The importance of some of these assumptions in driving model results was tested in sensitivity analyses.

Vasomotor symptoms

The model effectively measures the cost-effectiveness of an entire population using an estimated mean baseline rate of flushing across that population. In practice, the cost-effectiveness of treatment will also be influenced by treatment severity. The model assumes that vasomotor symptoms remain constant for the five year duration of the model. The NMA is used to estimate a probabilistic relative treatment effect for each intervention which, using baseline data, provided an estimate of the number of flushes per day with and without treatment. The model does not explicitly model a relationship between vasomotor symptoms and discontinuation rates although it would be expected that the effect of discontinuation on

vasomotor symptoms would be captured in the time period covered by the network meta-analysis. No costs were assigned to vasomotor symptoms but health state utilities were derived based on the number of daily flushes (see Appendix L.1.3.10).

Vaginal bleeding

Vaginal bleeding, a recognised side effect of combined HRT, was modelled as a dichotomous outcome. The network meta-analysis provided a relative treatment effect for each intervention which was converted into an absolute probability by applying these relative treatment effects to a probabilistic baseline probability. The model assumes that bleeding remains constant for the five year duration of the model. As with vasomotor symptoms no relationship between discontinuation and bleeding is explicitly modelled though again the network meta-analysis on bleeding is likely to capture the effects of discontinuation on bleeding for the time period covered by the network. A one-off cost is applied to this outcome (see Appendix L.1.3.7) which in the model is captured as a weighted average of those experiencing bleeding and those not experience bleeding, which incurs no cost for this outcome. A health state utility loss is also applied to the proportion who experience bleeding. Bleeding is not included as an outcome in the modelling of the population of women without a uterus.

Discontinuation of treatment

Beyond the six-month period covered by the NMA there are a range of assumptions made with respect to discontinuation over the 5-year timeframe of the model which are reported as sensitivity analyses. In the base case it is assumed that there is no further discontinuation after the first two cycles which is similar to the assumption made in the network meta-analysis which assumes no further discontinuation if the final measurement of discontinuation in the study is made before six months. There was some suggestion from studies that were included in the NMA which showed a levelling off in discontinuation if they included multiple time points. On the other hand there is observational data based on longer term follow-up which suggests, not surprisingly, that HRT use does decline further over time. The effects of discontinuation in the model largely impact on the costs of treatment. However, for patients on oral HRT it also impacts on the cycle risk of VTE, with those discontinuing reverting to baseline risk. Similarly for those on any HRT the risk of breast cancer is affected by the proportion continuing treatment who have a risk associated with HRT and the proportion who have discontinued who have a baseline risk.

Breast cancer

Breast cancer is a complex disease and it would technically be feasible to develop the model in such a way that reflected this complexity and the natural history of the disease. Such a model would require data on adverse events with breast cancer, the costs of those adverse events, their relative frequency, the relative frequencies of different stages of cancer on diagnosis, timing to events, and a typical treatment pathway for different stages over time along with any treatment complications and their frequency for example. However, it was not felt possible to develop such an approach to breast cancer outcomes given the timelines available to develop this guideline.

All models necessarily simplify the “real world” whilst aiming to reflect the essential and important features. The primary purpose of including breast cancer within this model is to acknowledge that it will have a differential impact on the “downstream” costs of interventions and a differential impact on morbidity and mortality, measured in the model by QALYs. The model uses UK published data and a relative treatment risk, where relevant, to estimate a per cycle incidence of breast cancer. The “downstream” cost of breast cancer is calculated as a weighted average of the cost of a breast cancer case and new cases across the timeframe of the model. Breast cancer has an associated mortality and this is modelled using a published estimate relevant to a UK context (see Appendix L.1.3.10). Rather than

modelling mortality as part of a more complex natural history the model makes the simplifying assumption that mortality in new cases occurs in the same cycle. The QALY loss from mortality is based on the woman's age at death and the remaining life expectancy of women of that age (this is explained in more detail in Appendix L.1.3.10). A utility loss, derived from the literature, is used to reflect the morbidity of breast cancer in surviving cases. A simplifying assumption was made that this morbidity would have a duration of 5 years. This was used because of the importance attached to 5-year survival in breast cancer but is to some extent arbitrary. However, the sensitivity of the model to the costs of breast cancer and QALY loss arising from it can be easily assessed.

VTE

VTE is a risk of oral (tablet) oestrogen, oestrogen and progestogen HRT or tibolone. Like breast cancer, there will be a differential impact on the "downstream" costs and QALYs depending on the intervention. The modelling approach is similar to that used in breast cancer and is a simplification of the natural history of VTE. The model estimates a per cycle incidence of new cases adjusted by the relative risk of treatment for oral HRT. A cost is then assigned to these cases which is the weighted average of the cost of a case and the number of cases. The model assumes that a certain proportion of cases have deep vein thrombosis and that the remainder have a pulmonary embolism. It is assumed that pulmonary embolism has a fatality rate and as with breast cancer it is assumed that mortality in a new case occurs in the same model cycle. As with breast cancer the QALY loss from mortality is determined by age at death and remaining life expectancy at that age. A utility loss is applied to those surviving with VTE which is assumed to last for 5 years. Sensitivity analysis was used to assess the extent to which the models results were affected by changes in assumptions about the cost and QALY loss associated with VTE.

L.1.3.7 Costs

Costs were based on an NHS and Personal Social Services perspective as outlined in the NICE reference case (The guidelines manual, NICE November 2012). The model has a duration of 5 years and therefore future costs were discounted at a rate of 3.5% in the base case analyses. The price year for costs was 2015.

Treatment costs

For the network meta-analysis that provides the estimates of treatment effectiveness, the GDG agreed that it was reasonable to construct the network by treatment class and route of delivery as opposed to individual drugs. The rationale for this being that they expected there to be little variation in treatment effectiveness by class. Where treatments are assumed to be equally effective it follows that the most cost effective treatment among them will be the cheapest. Therefore, the GDG agreed that the costing for each treatment class would be based on the cheapest drug in class that they would be willing to recommend if this drug was shown to be cost-effective. The costs used for treatment costs are shown in and are the costs for each three months cycle of 91 days unless otherwise noted in a table footnote. Where possible the costs of herbal therapies was based on products having a traditional herbal registration (THR) from the Medicines & Healthcare products Regulatory Agency (MHRA) and an indication for menopause. Products with THR are indicated in **Table 6**.

Table 6: Treatment costs

Class	Cost	Source
Oestrogen only oral	£5.49	Estradiol 2mg £5.07 for 84 tablets BNF March 2015

Class	Cost	Source
Oestrogen only patch	£12.61	Self-adhesive estradiol '50' patch £3.88 for a pack of 8 2 patches per week BNF February 2015
Oestrogen and progestogen oral	£9.97	Estradiol 1 or 2 mg 28 tablets plus norethisterone acetate 1 mg 12 tablets 3 x 28 tablet pack £9.20 (cyclical HRT)3 x 28 days treatment cycle (84 days) BNF March 2015
Oestrogen and progesterone patch	£36.04	Evorel Sequi £11.09 for a pack of 8 2 patches per week BNF March 2015
Tibolone	£33.67	Tibolone 2.5mg tablets £10.36 for 28 tablets NHS Drugs Tariff February 2015
Multibotanicals	£25.92	Vitabiotics Menopace Plus Tablets £15.95 for 56 tablets (http://www.hollandandbarrett.com[a])
SSRIs/SSNRIs	£3.75	Venlafaxine 37.5mg tablets £2.31 for 56 tablets NHS Drugs Tariff February 2015
Gabapentin	£3.78	Gabapentin 300mg tablets £4.15 for 100 tablets NHS Drugs Tariff February 2015
Isoflavones/Genistein/Soy	£17.09	Soya Isoflavones - 50mg £16.90 for 90 Tablets (Amazon.co.uk[b])
St John's Wort	£25.78	St John's wort extract 300 mg £8.50 for 50 tablets (http://www. schwabepharma.co.uk[c]) THR product
Black cohosh	£24.24	black cohosh extract 4.5 mg £7.99 for 30 tablets (http://www. chemistdirect.co.uk[d]) THR product
Valerian Root	£23.45	Dry extract from Valerian root (Valeriana officinalis L.) (equivalent to 135 to 167mg of Valerian root) 6 times per day £8.59 for 200 tablets (http://www. kalmsstress.com[e]) THR product
Acupuncture	£545	£65 initial appointment £40 per weekly session (for 12 weeks)[f] (http://www.ukacupuncture.co.uk[g])
Chinese herbal medicine	£112.11	Herbal Tablets Zhi Bai Di Huang Pian 2 tablets twice daily £9.24 for 30 tablets (http://www.ebay.co.uk[h])

(a) Full url: http://www.hollandandbarrett.com/shop/product/vitabiotics-menopace-plus-tablets-60083271 (accessed 1 March 2015)

(b) Full url: http://www.amazon.co.uk/Natures-Aid-Soya-Isoflavones-Genistein/dp/B000GY76L6 (accessed 27 February 2015)
(c) Full url: http://www.schwabepharma.co.uk/pages/products/menomood.php(accessed 17 March 2015)
(d) Full url: http://www.chemistdirect.co.uk/kira-menopause-relief/prd-ian (accessed 17 March 2015)
(e) Full url: http://www.kalmsstress.com/buyonline.html (accessed 17 March 2015)
(f) It is assumed that acupuncture is for 13 weeks (one cycle only) and that no further treatment costs are incurred in subsequent cycles
(g) Full url: http://www.ukacupuncture.co.uk/prices.php (accessed 1 March 2015)
(h) Full url: http://www.ebay.co.uk/itm/Zhi-Bai-Di-Huang-Pian-Wan-Menopause-Syndrome-Cystitis-Hot-Flushes-Osteoporosis-/141098048795?pt=LH_DefaultDomain_3&hash=item20da195d1b (accessed 8 March 2015)

In addition to the cycle costs of treatments an initial appointment with a GP to prescribe HRT was also included where applicable (**Table 7**).

Table 7: Costs of initial appointment

Parameter	Cost	Source
GP visit	£46	Per patient contact lasting 11.7mins (Netten 2014)

Costs relating to adverse outcomes

In addition to the costs of treatment there are costs related to bleeding, treatment of breast cancer and treatment of VTE. The costs relating to bleeding used in the base case analysis are shown in. It is assumed that if a woman experiences bleeding on treatment that this persists and will be investigated. This investigation is assumed to consist of an initial appointment with a gynaecologist and a transvaginal ultrasound scan with biopsy. It is additionally assumed that in a proportion of women that a diagnostic hysteroscopy will be necessary and in a small minority of these the procedure will need to be performed as a day case. Finally, it is assumed that there is a single appointment with a GP for follow-up. The proportion of women receiving the various components of this investigation is given in **Table 9**.

Thus women with bleeding incur a mean cost of £486 although this is a probabilistic parameter in the model which is sampled from a distribution estimated using the mean cost and upper and lower quartile range from NHS Reference Costs (see the Diabetes in Pregnancy update 2015 guideline for a fuller description of this method).

Table 8: Costs related to diagnosis of bleeding

Parameter	Mean cost	Lower Quartile	Upper Quartile
Initial gynaecologist appointment [a]	£151	£118	£188
Transvaginal ultrasound with biopsy [b]	£218	£136	£262
Outpatient diagnostic hysteroscopy with biopsy[c]	£211	£167	£255
Day case diagnostic hysteroscopy with biopsyd	£898	£748	£1,022

(a) For GP cost see Table 7
(b) NHS Reference Costs 2013/14, MA37Z Outpatient procedure transvaginal ultrasound with biopsy. Gynaecology
(c) NHS Reference Costs 2013/14, MA32Z Outpatient procedure Diagnostic hysteroscopy with biopsy. Gynaecology
(d) NHS Reference Costs 2013/14, MA32Z Day case procedure Diagnostic hysteroscopy with biopsy.

Table 9: Bleeding related probabilities

Item	% Patients	Notes
Appointment with gynaecologist	100%	Model assumption
Transvaginal ultrasound scan (TVUS) with biopsy	100%	Model assumption

Item	% Patients	Notes
Outpatient diagnostic hysteroscopy with biopsy	25%	GDG estimate of those who 'fail' TVUS
Day case diagnostic hysteroscopy with biopsy	2%	GDG estimate that 8% 'fail' with outpatient diagnostic hysteroscopy
GP follow-up appointment	100%	Model assumption

The costs of breast cancer were estimated using a report published by the Department of Health. This report broke the costs of breast cancer down by prognosis as shown in **Table 10** and the number of cases, which enabled a weighted mean cost of £9,501 to be calculated. This was then converted to a 2013 price year using the hospital and community health services (HCHS) index which gives a measure of combined pay and price inflation in healthcare. This index stood at 267 in 2008/09 and 290.5 in 2013/14, the latest year for which figures are currently available. This suggests an increase in prices of 8.8% should be applied to the cost derived from the Department of health. Therefore, the model assumed that each breast cancer case incurred costs of £10,337. This was a deterministic input parameter within the model but can be varied as part of a sensitivity analysis.

Table 10: Breast cancer costs by prognosis (2009)

Prognosis	n[a]	Cost
Excellent	7,519	£8,767
Good	5,326	£9,945
Moderate	1,253	£11,098
Unknown	313	£13,173

(a) Number of cases when screened – 1st year

In estimating the cost of VTE the model assumed that 94% of cases would have deep vein thrombosis (DVT) and that the remaining 6% of cases would be a pulmonary embolism (PE). This was taken from the NICE guideline on Venous Thromboembolism (CG92) and was based on the baseline risk with prophylaxis in a general medical population (13.4% DVT, 0.9% PE).

An average diagnosis and treatment cost of symptomatic DVT and non-fatal symptomatic PE was calculated using the data in **Table 11** and **Table 12** respectively. The item categories reflect those used in the NICE guideline on Venous Thromboembolism (CG92).

Table 11: Resource use and unit costs for symptomatic DVT

Item	Quantity	Unit cost	Total Cost	Notes
Diagnosis				
Doppler ultrasound	1	£62	£62	Based on 2013/14 NHS Reference Costs[a]
D-Dimer	1	£15.75	£15.75	Based on http://www.bivda.co.uk[b] and updated to 2014 prices using the HCHS index
Emergency admission	1	£97	£97	Based on 2013/14 NHS Reference Costs[c]
Treatment				
LMWH	7	£8.47	£59.29	Fragmin, BNF January 2015 15,000 units daily

Item	Quantity	Unit cost	Total Cost	Notes
Full blood count	2	£6.68	£13.35	2006/07 unit cost = £5.74 Updated to 2014 prices using HCHS index
Warfarin	119.52	£0.04	£4.48	Assumed 5mg daily 3 months x 69% (distal) 6 months x 31% (proximal)[d] 28 tablet pack = £1.05, BNF January 2015
Hospital stay	0.4 days	£589	£235.60	10% x 4 daysd 2006/07 unit cost = £192/day Updated to 2014 prices using HCHS index
LMWH instruction	0.45 hours	£100	£45	90% x 30 minutes Unit Costs of Health and Social Care 2014[e]
Graduated compression stockings	6	£0.50	£2.93	Over 2 years, 2nd year discounted at 3.5% NHS Drugs Tariff 2014
Anticoagulation clinics	9.93 visits	£56	£243.53	Distal = 9 visits Proximal = 12 visitsd £56 first visit £21 follow-up Based on 2013/14 NHS Reference Costs[f]
Ambulance transport to anticoagulation clinic	0.1	£231	£23.10	10% of visits Based on 2013/14 NHS Reference Costs[g]
Total				**£802.03**

(a) RA24Z, outpatient
(b) Full url: http://www.bivda.co.uk/Portals/0/Documents/Technology%20Briefs/Template%2015.pdf (accessed 1 March 2015)
(c) VB11Z, Type 01, Emergency Medicine, No Investigation with No Significant Treatment
(d) NICE guideline on Venous Thromboembolism (CG92)
(e) Based on hospital Band 5 nurse and per hour of patient contact
(f) WF01B, Anticoagulant service
(g) ASS02

Table 12: Resource use and unit costs for non-fatal symptomatic PE

Item	Quantity	Unit cost	Total Cost	Notes
Diagnosis				
CT pulmonary angiogram	1	£124	£124	Based on 2013/14 NHS Reference Costs[a]
Chest x-ray	1	£33.72	£33.72	2006/07 unit cost = £29 Updated to 2014 prices using HCHS index
ECG	1	£31.40	£31.40	2006/07 unit cost = £27 Updated to 2014 prices using HCHS index
D-Dimer	1	£15.75	£15.75	See Table 11
Emergency admission	1	£97	£97	See Table 11

Item	Quantity	Unit cost	Total Cost	Notes
Treatment				
LMWH	7	£8.47	£59.29	See Table 11
Full blood count	2	£6.68	£13.35	See Table 11
Warfarin	183	£0.04	£6.86	See Table 11 6 months
Hospital stay	5.4 days	£589	£235.60	See Table 11 90% x 6 days
LMWH instruction	0.05 hours	£100	£5	See Table 11
ICU stay	0.7 days	£1,057	£739.90	Based on 2013/14 NHS Reference Costs[b] 10% x 7 days
Graduated compression stockings	6	£0.50	£2.93	See Table 11
Anticoagulation clinics	12 visits	£56	£287	See Table 11
Ambulance transport to anticoagulation clinic	0.1	£231	£23.10	See Table 11
Total				**£4,619.91**

(a) RA10Z, outpatient
(b) XC07Z cardiac adult surgical patients predominate

In line with the NICE guideline on Venous Thromboembolism (CG92) the cost of post-thrombotic syndrome (PTS) and chronic thromboembolic pulmonary hypertension (CTPH) was also included to derive a weighted mean cost of VTE. These costs are given in Table 13 and Table 14 shows how the weighted mean cost of VTE is calculated using the probability of various VTE associated events.

Table 13: Other treatment related costs

Event	Cost	Notes
Post-thrombotic syndrome	£8,782	NICE CG92[a]
Chronic thromboembolic pulmonary hypertension	£80,385	NICE CG92[a]

(a) Updated to 2014 prices using HCHS index

Table 14: Weighted mean cost of VTE[a]

Event	Probability	Weight	Unit cost	Weighted Cost	Notes
VTE	100%	-	-	-	
DVT	93.7%	-	-	-	
PE	6.3%	-	-	-	
Asymptomatic DVT	93.8%	0.938 x 0.94 = 0.879	£0.00	£0.00	
Symptomatic DVT	6.2%	0.062 x 0.94 = 0.058	£802.03	£46.60	See Table 11
Non-fatal PE	94%	0.94 x 0.06 = 0.059	£4,619.91	£273.32	See Table 12
PTS after symptomatic DVT	25%	0.25 x 0.062 = 0.015	£8,782	£127.55	See Table 13

Event	Probability	Weight	Unit cost	Weighted Cost	Notes
PTS after asymptomatic DVT	15%	0.15 x 0.938 = 0.132	£8,782	£1,157.81	See Table 13
CTPH after non-fatal PE	0.75%	0.0075 x 0.059 = 0.0004	£80,385	£35.67	See Table 13
Weighted Mean cost				**£1,641**	

(a) Probabilities taken from the NICE guideline on Venous Thromboembolism (CG92)

L.1.3.8 Baseline data

The NMA generates a measure of treatment effect for each drug class relative to placebo. It was assumed that placebo could be used to represent a no treatment option. However, the event rate in the placebo arm (for discontinuation and bleeding) and mean number of hot flushes per day (for vasomotor symptoms) varies considerably from trial to trial and may not necessarily reflect the current baseline risk in England and Wales. A similar issue presents itself with respect to the VTE and breast cancer risks estimated from conventional pair-wise meta-analysis.

Therefore, alternative sources of baseline data were sought from the published literature. The model inputs for baseline data are shown in Table 15 together with the probabilistic parameters that define the distribution used in the probabilistic sensitivity analysis undertaken to reflect the inherent uncertainty in these point estimates of baseline risk.

Table 15: Model baseline risks

Outcome	Risk/mean	Alpha	Beta	Distribution	Notes
Discontinuation	-	-	-	N/A	Placebo arm in trial[a,b]
Bleeding	0.107	29	242	Beta	Astrup 2004
Breast cancer[c]	0.0028	5,608	2,016,751	Beta	Cancer Research UK
VTE	0.0042	158	37,730	Beta	RCOG 2011
Hot flushes	3.0 flushes/day	0.568	5.280	Gamma	Brown 2009

(a) The model used the sampled discontinued rates for placebo from the network meta-analysis
(b) No treatment cannot be discontinued and therefore to derive the absolute risk from the relative treatment effect derived from the NMA is was necessary to use the discontinuation rates from the placebo arms of trials for studies included in the network meta-analysis
(c) This is the incidence rate or risk per annum in women aged 50-54

L.1.3.9 Treatment effectiveness

Appendix L.1.3.8 outlines the baseline risk of the model outcomes. For the outcomes in the network meta-analysis the absolute probability for each treatment of discontinuation or bleeding or the mean rate of hot flushes per day in any simulation is derived by applying the relative treatment effect from a network-meta analysis iteration for that treatment to the sampled baseline risk in that simulation.

In order to estimate the absolute risk of breast cancer and VTE, a relative risk derived from the pairwise meta-analysis undertaken for the guideline is applied to the baseline risk for those outcomes (see Table 15). The relative risks for those outcomes is shown in Table 16.

Table 16: Relative risk for breast cancer and VTE on HRT[a]

Event	Relative Risk	SElogRR	Distribution	Notes

Event	Relative Risk	SElogRR	Distribution	Notes
Breast cancer	1.49	0.144	Log-normal	Pairwise meta-analysis
VTE	2.12	0.437	Log-normal	Pairwise meta-analysis

(a) It is assumed that the risk of VTE only applies to oral HRT

L.1.3.10 Health state utilities and Quality Adjusted Life Years (QALYS)

The health state utilities used in the model are shown in Table 17. Details of how they were derived is described below. The model assumes that the decision maker has a willingness to pay for a QALY.

Table 17: Model health state utilities

Event	Health State utility decrement	Notes
Death	0.82	Based on population norm[a]
Hot flush	0.021	See Appendix L.1.3.10**Error! Reference source not found.**
Bleeding	0.01	Daly 1993
Breast cancer	0.28	See Appendix L.1.3.10 Peasgood 2010[b]
Venous thromboembolism	0.007	See Appendix L.1.3.10

(a) EQ5D instrument completed by 3,395 people resident in the UK
(b) This was based on the higher range of an estimate for metastatic breast cancer

An annual discount rate of 3.5% was used to assess changes in health state utility over time. For many model events the QALY loss is determined by the 5-year duration of the model. However, events leading to death have a far longer lasting impact given the typical life expectancy of the model population. In order to calculate the QALY loss of a death the model assumes that a 0.82 health state utility decrement is experienced over the remaining life expectancy of women of that age. This is illustrated graphically in Figure 2 and Table 18 shows the model life expectancies, which reflect that the model assumed that women start treatment at 50 years old but will continue on treatment for 5 years.

Figure 2: QALY loss from death

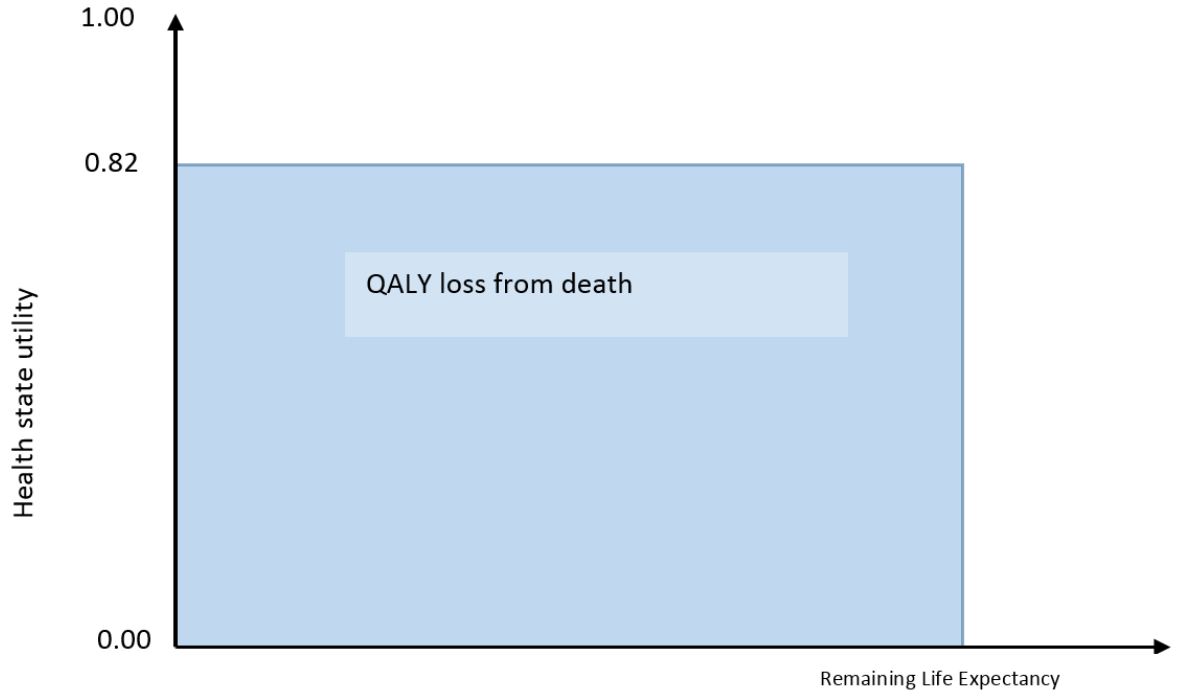


Table 18: Woman's remaining life expectancy by age[a]

Age	Remaining life expectancy
50 years old	34 years
51 years old	33 years
52 years old	32 years
53 years old	32 years
54 years old	31 years

(a) National Life Tables, England & Wales, 1980-82 to 2011-13, ONS 2013

Vasomotor symptoms

The frequency of hot flushes is often used as a proxy for the severity of menopausal symptoms. Therefore a reduction in the frequency of flushing can be expected to bring a relief in symptoms and concomitant improvement in health related quality of life. Based on GDG opinion, the model assumed that each unit reduction in mean hot flushes per day would give the same health state utility gain irrespective of the severity at baseline. In other words the gain in health state utility in going from 1 hot flush to 0 hot flush per day is the same as going from 10 hot flushes to 9 hot flushes per day.

The model relationship between mean hot flushes per day and health state utility was estimated using two time trade-off studies (Daly 1993, Zethraeus 1997) on women's preferences with respect to menopausal symptoms. This represented a departure from the NICE reference case method of eliciting utilities, but was considered to be the best available data from the literature on preferences related to menopausal symptoms, especially given the need to link changes in health state utility to the outcome used in network meta-analysis to assess treatment efficacy.

In the base case analysis the GDG agreed that 20 flushes per day would be reasonable as an upper limit of severe symptoms, whilst acknowledging that at the very extreme two flushes per hour are possible. In the British study (Daly 1993) severe menopausal symptoms were estimated to have a health state utility of 0.64. Therefore, assuming a linear relationship

and perfect health in the absence of any flushing, the health state utility decrement per hot flush was calculated as follows:

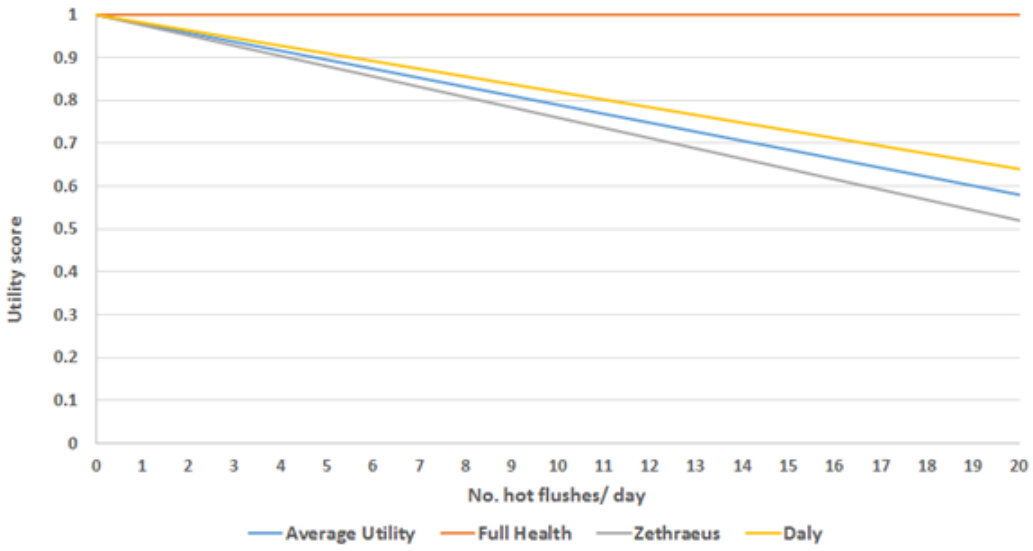
Health state utility decrement per hot flush = (1.0 – 0.64) ÷ 20 = 0.018

In the Swedish study (Zethraeus 1997) severe menopausal symptoms were rated as having a health state utility of 0.52, from which the utility decrement per hot flush was calculated.

Health state utility decrement per hot flush = (1.0 – 0.52) ÷ 20 = 0.024

The model used the average derived from these two studies as the base decrement in health state utility per hot flush and the relationship between health state utility and hot flushes per day is presented in Figure 3.

Figure 3: Illustration of model relationship between health state utility and mean number of hot flushes per day



It was assumed that the treatment effect was constant through all cycles of the model and therefore the overall QALY gain from a reduction in hot flushes would be the discounted gain in health state utility compared to no treatment multiplied by the 5-year duration of the model.

Venous thromboembolism

There is a per cycle risk of VTE and 6% of VTE cases are assumed to have a pulmonary embolism. In those having a pulmonary embolism it is assumed there is a mortality rate of 6%, which in the model is assumed to occur in the same cycle as the pulmonary embolism event. So the overall mortality risk of VTE is given by:

Mortality rate of VTE = 0.063 x 0.06 = 0.0038

Figure 2 shows graphically how the model calculates the QALY loss associated with a death arising from a pulmonary embolism.

A health state utility decrement for VTE not resulting in death was estimated by applying a health state utility to a range of VTE related events weighted by their probability (see also Appendix L.1.3.7). The events, their weights and utilities are shown in Table 19.

Table 19: Health state utilities associated with VTE events and their weight

Event	Weight	Health state utility decrement	Weighted health state utility
Surviving VTE[a]	1		
Asymptomatic DVT no PTS	0.750	0.000	0.0000
Asymptomatic DVT with PTS	0.132	0.018	0.0024
Symptomatic DVT no PTS	0.043	0.000	0.0000
Symptomatic DVT with PTS	0.015	0.018	0.0003
PE with no CTPH	0.059	0.060	0.0035
PE with CTPH	0.0004	0.235	0.0001
Weighted mean VTE health state utility decrement			0.0063

(a) The weights in this Table are related to those in Table 13 but are based on those surviving VTE rather than all cases of VTE and hence are higher by a factor of 1 ÷ 0.9962

The model assumes that in cases that experience a VTE that the associated health state utility loss will persist for the remaining model cycles. This is likely to over-estimate the QALY loss associated with VTE if treatment is able to quickly return them to a pre-VTE health state. However, this needs to be interpreted in the light of the low baseline risk and the low health state utility loss associated with a surviving VTE. Therefore, even assuming this remaining model cycle duration health state utility loss, the absolute QALY loss remains very small. The impact of changing the QALY loss associated with VTE on the model's results was tested with sensitivity analysis.

Breast cancer

For those surviving breast cancer it is assumed in the base case that breast cancer effects their health state utility for a period of 5 years after becoming an incident case. The QALY loss in a breast cancer survivor is shown by the blue shaded area inFigure 4.

In addition there is also a QALY loss associated with breast cancer mortality. The model assumed a breast cancer mortality rate of 15% (ONS, 2011). As noted earlier a simplifying assumption was made that death occurs in the same cycle as the incident case, with the QALY loss being calculated using the approach illustrated in Figure 2. A more sophisticated model could have used a time to event approach in which case the QALY loss with breast cancer mortality would more closely resemble the light shaded blue area in Figure 5, with the darker shaded error depicting declining health related quality of life over a period of time prior to death.

Thus the simplified approach used in this model will tend to over-state the QALY loss from breast cancer mortality by assuming immediate death. However, this just contributes to the overall QALY loss from all breast cancers and again the impact of QALY loss from breast cancer was assessed through sensitivity analysis, which can be adjusted by altering the duration for which breast cancer causes a reduction in health related quality of life and/or the health state utility decrement attributable to breast cancer. Furthermore, it should be noted that the breast cancer mortality parameter in the model is based on survival rates at 5 years, in which case most of the QALY loss from mortality would be as a result of decreased longevity.

Figure 4: QALY loss from a breast cancer survivor

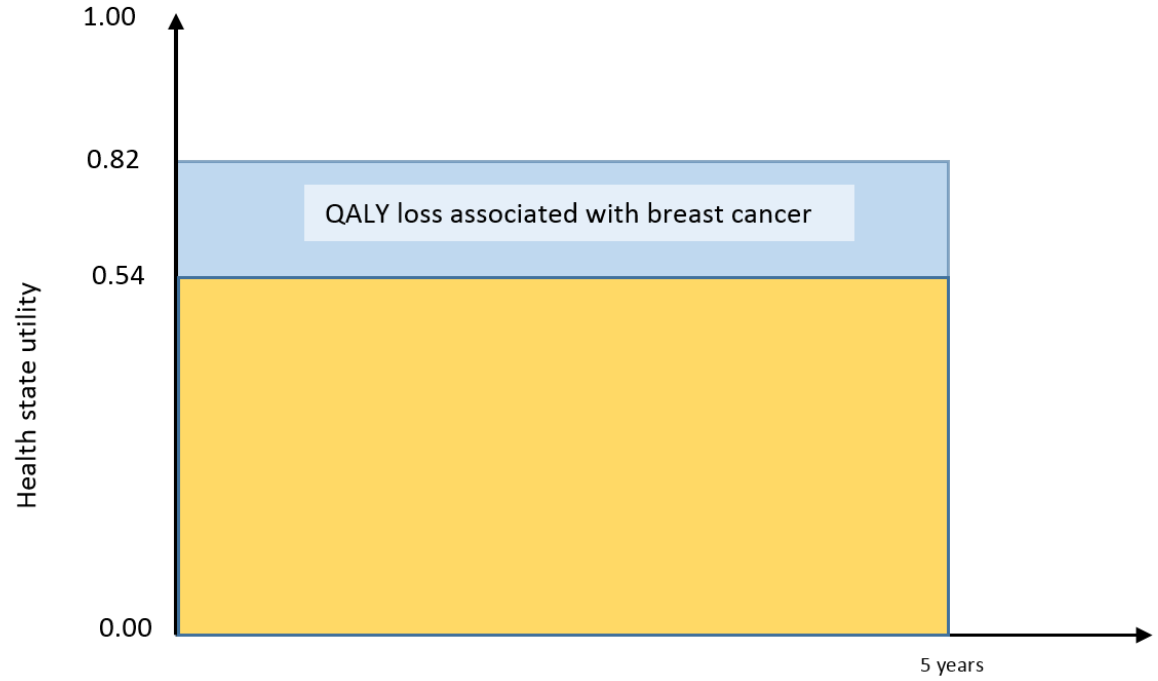
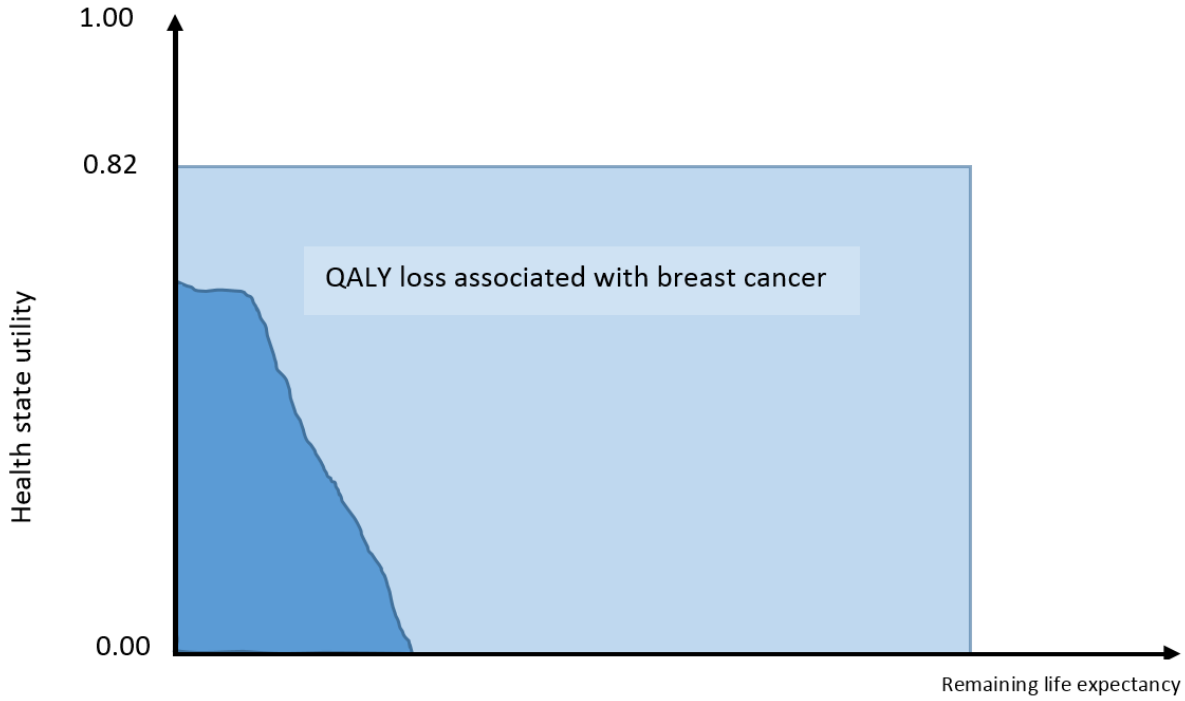


Figure 5: QALY loss associated with breast cancer mortality if using a time to event approach



L.1.3.11 Sensitivity analysis

All model analyses presented in Appendix L.1.4 are based on probabilistic sensitivity analysis to reflect uncertainty in parameter estimates. However, for some variables there is parameter uncertainty other than that accounted for by sampling variability. Therefore, a number of sensitivity analyses were undertaken whereby a deterministic input is changed before running the probabilistic sensitivity analysis. These can help assess how sensitive the model is to changes in particular parameters especially where simplifying assumptions were used. Furthermore, these sensitivity analyses can also be used to validate the model by checking that the model changes in a predictable way in response to its inputs.

L.1.4 Results

L.1.4.1 Women with a uterus

Base Case analysis

The results for the base case analysis are presented in Table 20 and Figure 6, Figure 7 and Figure 8. This suggests that black cohosh and non-oral oestradiol and progestogen can be considered the most cost effective treatment in this analysis based on the treatments with the highest net mean benefit. Non-oestradiol and progestogen has the highest level of QALY gain and with an incremental cost-effectiveness ratio (ICER) falling within a willingness to pay threshold of £20,000-£30,000 per QALY. The net mean benefit was calculated as follows:

Net mean benefit = Mean QALY x £20,000 – Mean Cost

The mean QALY is multiplied by £20,000 as that represent the decision makers willingness to pay for a QALY gain of that magnitude and the mean cost is then subtracted from it to give a measure of benefit less the cost of achieving that benefit. The mean costs and mean QALYs are based on the cost and QALY generated for each treatment in each simulation, over 10,000 simulations in this case. Treatments that are marked as “dominated” have unambiguously preferred treatment alternatives offering a higher mean QALY and lower mean costs.

Also presented is the probability that each treatment is cost-effective which reflects the degree of uncertainty in the results and is calculated from the number of times that a particular intervention is the most cost-effective over all the individual simulations. Whilst this is not independent of the mean costs and benefits it will also reflect uncertainty in the point estimates of relative treatment effects and intervention with a lower point estimate of effectiveness may have a relatively high probability of being cost-effective if it has particularly wide confidence intervals. ICERs are calculated relative to the next best non-dominated treatment alternative.

Table 20: Base case results for women with a uterus based on 10,000 simulations[a]

Treatment	Mean cost	Mean QALY	Net mean benefit	Probability cost-effective	ICER
No treatment	£0	0.0000	£0	2.3%	n/a
SSRIs/SMRIs	£34	0.0415	£797	18.6%	£813
Gabapentin	£52	0.0587	£1,122	14.9%	£1,042
Isoflavones/Genistein/Soy	£312	0.1089	£1,866	2.3%	Extended Dominance
Oestradiol + progestogen oral	£385	0.0784	£1,183	1.8%	Dominated

Treatment	Mean cost	Mean QALY	Net mean benefit	Probability cost-effective	ICER
Valerian Root	£437	0.0001	-£436	0.0%	Dominated
Black cohosh	£448	0.1646	£2,845	27.1%	£3,740
Multibotanicals	£483	0.0504	£524	5.0%	Dominated
Acupuncture	£545	0.1084	£1,624	6.4%	Dominated
Tibolone	£598	-0.0017	£1,019	0.0%	Dominated
Oestradiol + progestogen non- oral	£888	0.1845	£2,801	19.3%	£22,165
Chinese herbal medicine	£2,009	-0.0018	-£2,044	0.0%	Dominated

(a) Mean costs and mean QALYs are calculated relative to no treatment

Figure 6: Base case results for women with a uterus based on 10,000 simulations illustrated on the cost-effectiveness plane

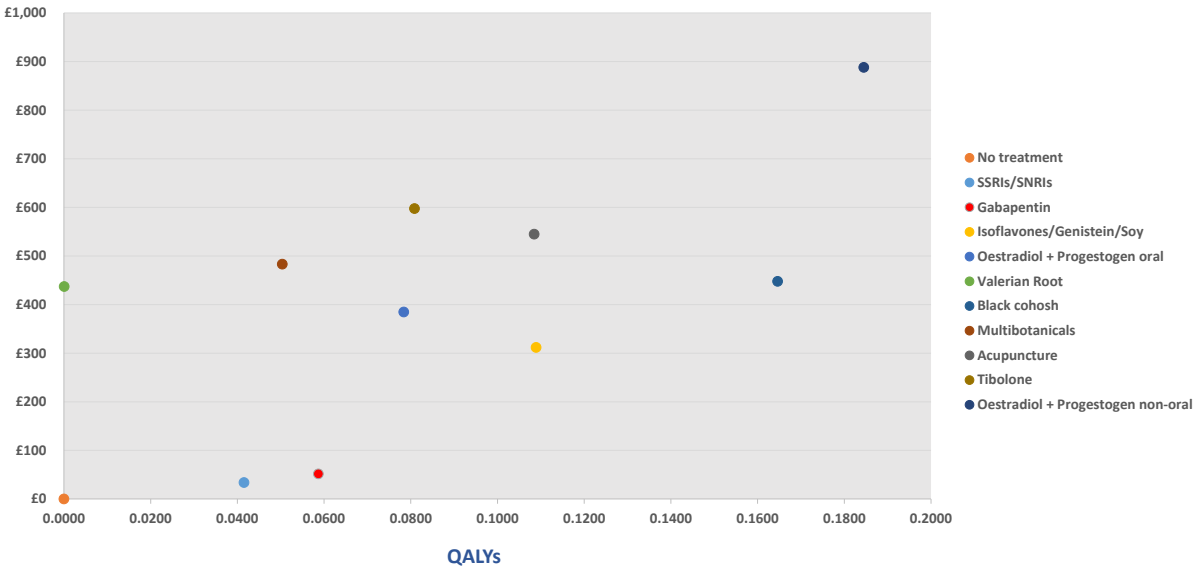


Figure 7: Cost-effectiveness plane for base case analysis for women with a uterus from 1,000 simulations

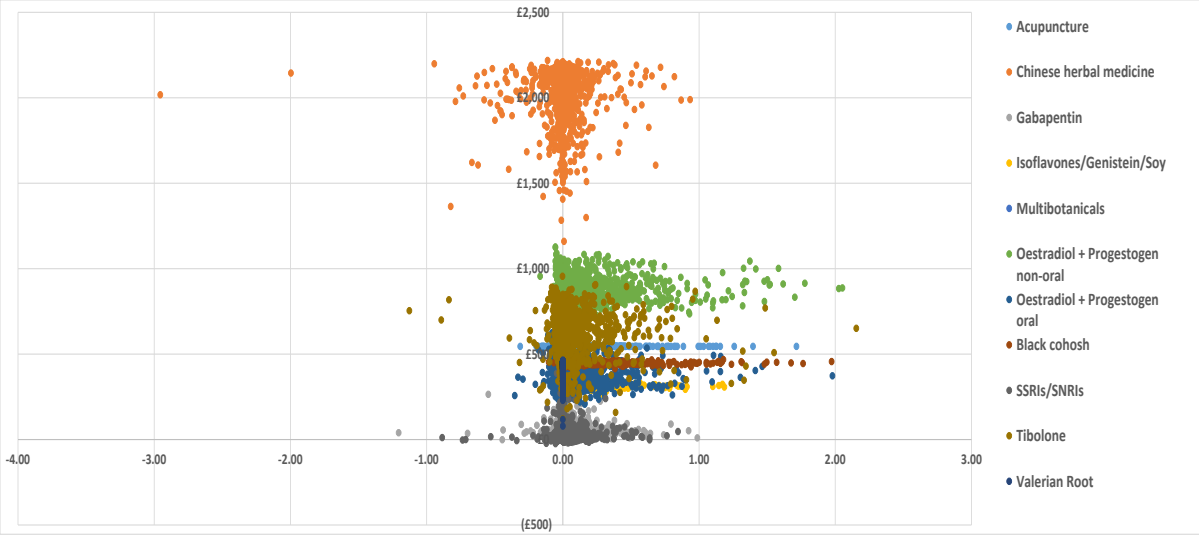
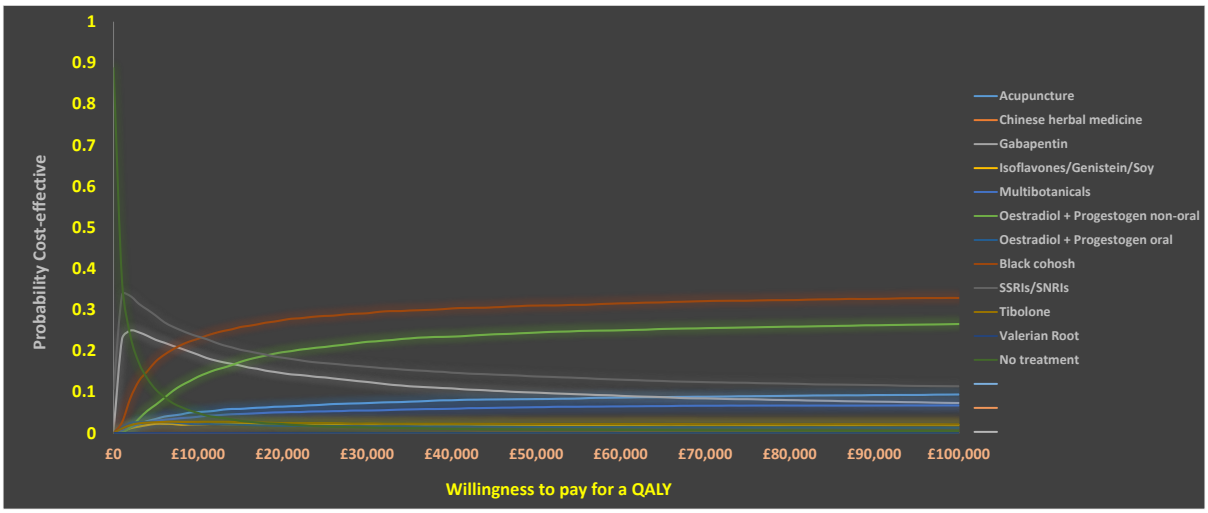


Figure 8 shows the cost-effectiveness acceptability curve (CEAC) which illustrates how the probability of a treatment being the most cost-effective varies with willingness to pay. At a zero willingness to pay the CEAC shows the probability of the treatment being the cheapest option. In this case no treatment has the highest probability of being the cheapest option, meaning that on average downstream savings from reduced adverse events are not sufficient to more than offset the treatment costs. At £20,000 per QALY the charted probabilities reflect the values in Table 20.

Figure 8: Cost-effectiveness acceptability curve of base case analysis in women with a uterus based on 10,000 simulations



Sensitivity analysis

i. Symptom severity

The base case analysis is based on a mean baseline rate of 3 flushes per day, albeit with sampling variability. It should also be noted that the distribution used for probabilistic sensitivity analysis is right skewed meaning that in most of the simulations the mean number of flushes per day would be less than 3 flushes per day. Three flushes per day would be considered as a mild/moderate level of flushing. Whilst there might be uncertainty around the mean rate of flushing across the entire population the presenting woman will often have a good knowledge of the number of flushes she is experiencing and therefore if the number of flushes per day is used as proxy for severity, then a severity based on flushes per day threshold could be used to determine treatment if cost-effectiveness varied by symptom severity.

In this sensitivity analysis the rate of hot flushes per day was varied whilst keeping all other model inputs at their base case value. Figure 9 show that the probability of non-oral oestradiol and progestogen being cost-effective increases with increasing severity. Figure 10 indicates how the mean net benefit of effective treatment is a function of increasing severity.

The CEAC and cost-effectiveness plane for a woman with a uterus having 10 hot flushes per day is shown in Figure 11 and Figure 12 respectively. This shows that non-oral oestradiol and progestogen has a high probability of being cost-effective at most willingness to pay thresholds.

Figure 9: Probability of different treatments being cost-effective varying the mean hot flushes per day in women with a uterus

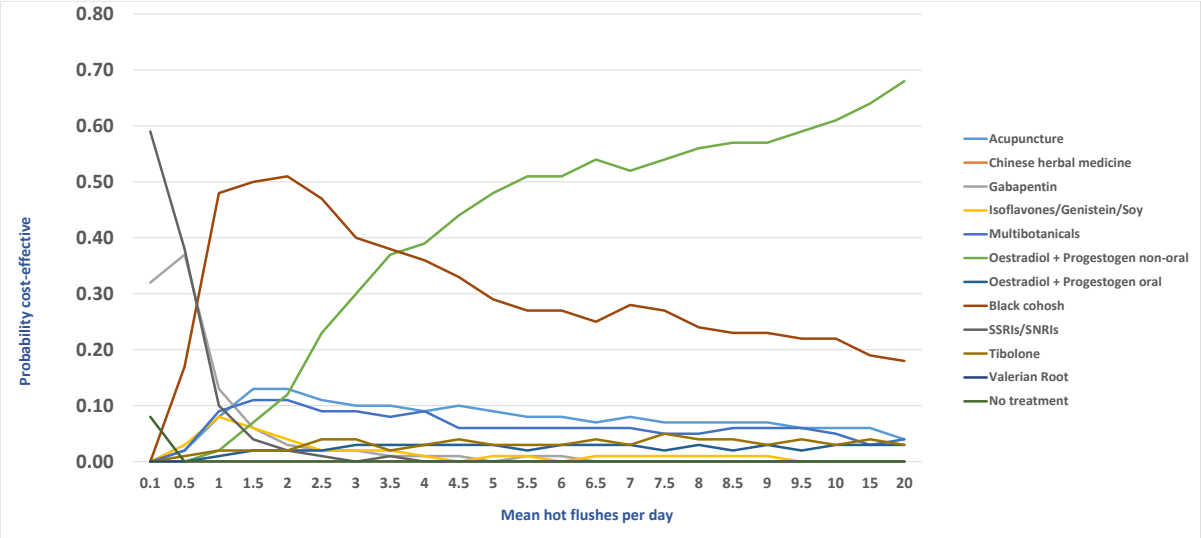


Figure 10: Mean net benefit varying the mean hot flushes per day

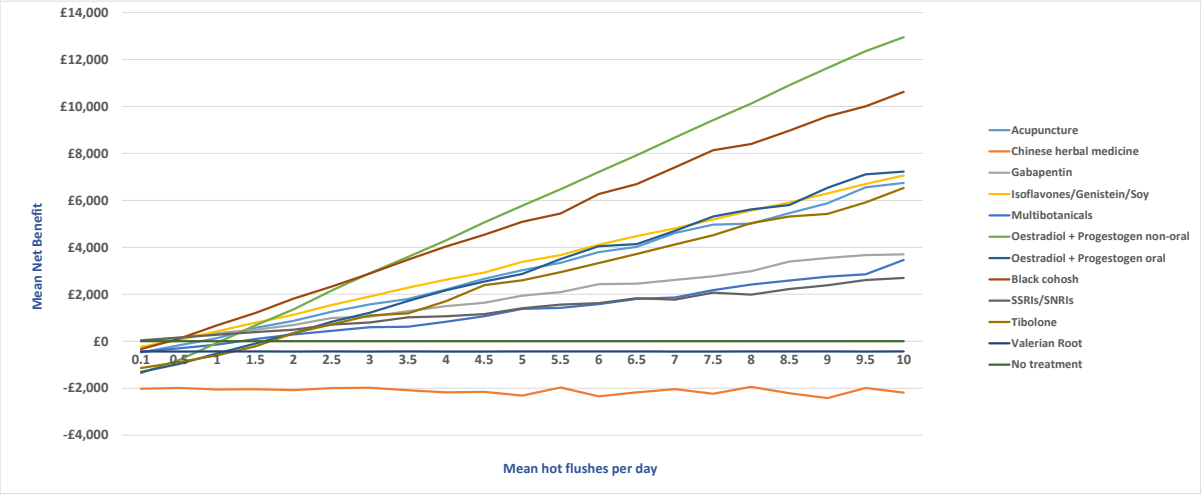


Figure 11: CEAC with a mean of 10 hot flushes per day

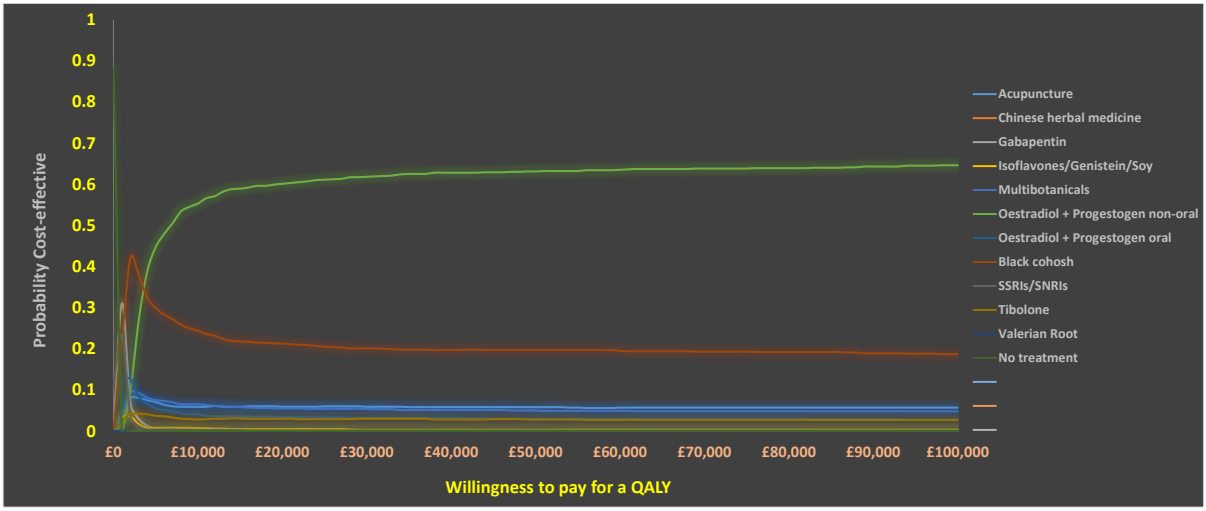
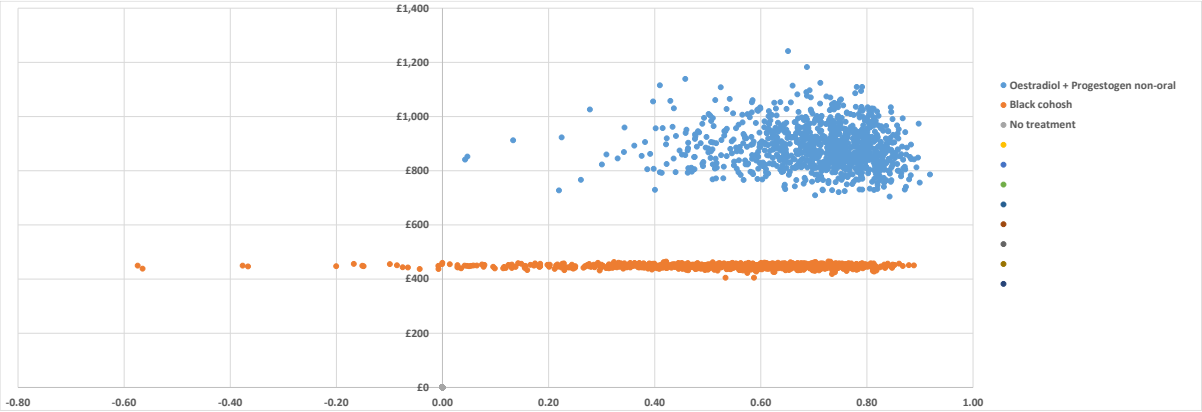


Figure 12: Cost-effectiveness plane showing the results of 1,000 simulations for non-oral oestrogen and progestogen and black cohosh



ii. Varying the costs and QALY losses from breast cancer

The model suggests that non-oral oestrogen and progestogen is cost-effective for the treatment of menopausal symptoms and increasingly so with increasing symptom severity, represented by an increasing number of mean hot flushes per day in this model. However, breast cancer is a potential adverse outcome of non-oral oestrogen and progestogen and it is important therefore to assess to what extent the model's results are driven by the assumption made with respect to the costs and QALY loss associated with breast cancer. Clearly, if the cost and/or QALY loss for breast cancer were over-estimated in the base case then the finding that non-oral oestrogen and progestogen was cost-effective would simply be strengthened. Therefore, it is more important to assess the implications for the model if the cost of breast cancer and/or QALY loss had been under-estimated and that is what has been undertaken in this sensitivity analysis.

First the implications of increasing the breast cancer costs to £30,000 were assessed whilst keeping other deterministic model parameters at their base case value and the results of this analysis are shown in Table 21. Whilst there is an increase in the costs of HRT and tibolone, as would be expected, compared to the base case results presented in Table 20, this has only a very minor impact on both the net mean benefit and the probability of a treatment being cost-effective. Non-oral oestrogen and progestogen remains borderline the most cost-effective treatment if using a £20,000 to £30,000 per QALY willingness to pay threshold. It should be remembered that this analysis is for a sampled population baseline rate of flushing with a mean of 3 hot flushes per day but with a majority of the simulations having less than the mean. If the analysis is repeated for a woman presenting with 5 mean hot flushes per day the probability that non-oral oestrogen and progestogen is cost-effective is 41% (compared to 48% when using the lower base case breast cancer costs, see Figure 9), with the probability of cost-effectiveness increasing with higher mean hot flushes per day.

Table 21: Results with breast cancer costs of £30,000 for women with a uterus based on 1,000 simulations[a]

Treatment	Mean cost	Mean QALY	Mean Net Benefit	Probability cost-effective	ICER
No treatment	£0	0.0000	£0	2.4%	n/a
SSRIs/SNRIs	£35	0.0369	£703	19/1%	Dominated
Gabapentin	£51	0.0633	£1,215	13.9%	£803
Isoflavones/Genistein/Soy	£312	0.1086	£1,860	1.6%	Extended dominance
Valerian Root	£432	0.0001	-£431	0.0%	Dominated

Treatment	Mean cost	Mean QALY	Mean Net Benefit	Probability cost-effective	ICER
Black cohosh	£448	0.1638	£2,829	27.6%	£3,949
Multibotanicals	£481	0.0563	£646	6.7%	Dominated
Oestradiol + progestogen oral	£512	0.0829	£1,146	1.7%	Dominated
Acupuncture	£545	0.1046	£1,547	5.8%	Dominated
Tibolone	£707	0.0755	£802	1.6%	Dominated
Oestradiol + progestogen non-oral	£1,015	0.1876	£2,737	19.6%	£23,865
Chinese herbal medicine	£2,011	0.0035	-£1,941	0.0%	Dominated

(a) Mean costs and mean QALYs are calculated relative to no treatment

Second, the implication of increasing the QALY loss associated with breast cancer was explored. The QALY loss is simply the health state utility loss multiplied by the duration of this loss (modified by the annual discount rate of 3.5%) and so either parameter can be altered. In this sensitivity analysis the duration of health state utility loss was increased to find the threshold at which the ICER for non-oral oestrogen and progestogen increased to over £30,000 per QALY.

For a sampled population having a mean of 3 hot flushes per day, increasing the duration of health state utility loss (assuming the health state utility remains constant) from 5 years to 7 years resulted in the ICER for non-oral oestrogen and progestogen, still the treatment with the highest QALY, increasing to £30,859 per QALY.

For a woman with a mean of 5 hot flushes per day the ICER for non-oral oestrogen and progestogen is well under a £20,000 per QALY threshold (£13,187) even if the duration of health state loss is increased to 34 years, the effective remaining life expectancy. Even if the duration is kept at 34 years and the costs of breast cancer almost tripled to £30,000 the ICER for non-oral oestrogen and progestogen for a woman with a mean of 5 hot flushes per day is £18,171 per QALY and would still be considered cost-effective at a willingness to pay threshold of £20,000 per QALY.

iii. Varying the costs and QALY losses from VTE

Women taking oral (tablet) HRT and tibolone in the model are assumed to have an increased risk of VTE. None of these treatment would be considered cost-effective based on the base case analysis and therefore if the VTE risk, VTE costs and/or QALY loss from VTE had been under-estimated then that would simply strengthen the finding of the base case analysis.

So in this sensitivity analysis, the impact of over-estimating these parameters was investigated by setting the relative risk for oral HRT and tibolone to 1, as this removes anything but stochastic differences between treatments in terms of VTE cases and any additional cost and QALY loss associated with VTE. The results of this analysis are shown in **Table 22**.

Table 22: Results with no increased risk of VTE for women with a uterus based on 1,000 simulations[a]

Treatment	Mean cost	Mean QALY	Mean Net Benefit	Probability cost-effective	ICER
No treatment	£0	0.0000	£0	2.2%	n/a
SSRIs/SNRIs	£33	0.0446	£860	18.6%	£735
Gabapentin	£53	0.0523	£994	15.0%	£2,580

Treatment	Mean cost	Mean QALY	Mean Net Benefit	Probability cost-effective	ICER
Isoflavones/Genistein/Soy	£311	0.1094	£1,878	2.1%	Extended dominance
Oestradiol + progestogen oral	£383	0.1034	£1,685	2.0%	Dominated
Valerian root	£435	0.0001	-£434	0.0$	Dominated
Black cohosh	£448	0.1613	£2,778	25.8%	£3,628
Multibotanicals	£481	0.0524	£566	5.3%	Dominated
Acupuncture	£545	0.1047	£1,549	7.5%	Dominated
Tibolone	£615	0.0974	£1,333	3.1%	Dominated
Oestradiol + progestogen non-oral	£887	0.1856	£2,825	18.4%	£18,083
Chinese herbal medicine	£2,030	0.0010	-£2,009	0.00%	Dominated

(a) Mean costs and mean QALYs are calculated relative to no treatment

As expected the mean cost of oral oestrogen and progestogen and tibolone falls and the mean QALY increases compared to the base case analysis results (see **Table 20**). However, this does not generate a big enough increase in mean net benefit to make much difference to the cost-effectiveness results or ranking.

iv. Varying treatment costs

Chinese herbal medicine has a very high cost relative to other herbal preparations. However, even if its cost is reduced to £3 from £112, holding all other deterministic inputs at their base case value, it still only has a 4% chance of being the most cost-effective treatment and its mean net benefit remains negative (-£29 based on 500 simulations). This is because it produces so little QALY gain (0.0001 based on 500 simulations) compared to no treatment.

If black cohosh was increased to the same cost as non-oral oestrogen and progestogen holding all other deterministic inputs at their base case value, then oestrogen and non-oral progestogen now has a slightly higher probability of being cost-effective than black cohosh (22% versus 21%) and a mean net benefit of £200 more than black cohosh. In the base case analysis black cohosh had a £44 higher mean net mean benefit when a QALY was valued at £20,000 per QALY. However, this is still a relatively small change even when in the majority of the simulations there is less than a mean of 3 flushes per day.

If the sensitivity analysis is run for a woman with severe vasomotor symptoms, a mean of 20 hot flushes per day, then even if the 3 month treatment cost of black cohosh is reduced to £5 from £24.24 in the base case analysis, then non-oral oestrogen and progestogen still has a 66% probability of being the most cost-effective compared to 20% for black cohosh.

v. Varying discontinuation assumptions

In the base case analysis the model assumes that no further discontinuation occurs after the first two cycles but the model is constructed to allow a number of alternative assumptions.

a) Discontinuation at the same rate as the first two cycles for the remaining 18 cycles

Holding all other inputs constant at their base case value, this sensitivity analysis has a negligible impact on the relative cost-effectiveness of black cohosh and non-oral oestrogen and progestogen, with them both having an almost identical mean net benefit.

b) 20% have discontinued by the end of 5 years

Again holding all other inputs constant at their base case value, this sensitivity analysis has a negligible impact on the relative cost-effectiveness of black cohosh and non-oral oestrogen and progestogen, the two most cost-effective treatments, with them both having an almost identical mean net benefit.

c) 40% have discontinued by the end of 5 years

This has the impact of producing a small increase in the cost-effectiveness of non-oral oestrogen and progestogen when compared to black cohosh, with the ICER falling to £16,304 per QALY.

d) 80% have discontinued by the end of 5 years

Making this assumption about discontinuation leads to the largest impact on the relative cost-effectiveness between oestrogen and progestogen when compared to black cohosh. The ICER for oral oestrogen and progestogen is now £8,054

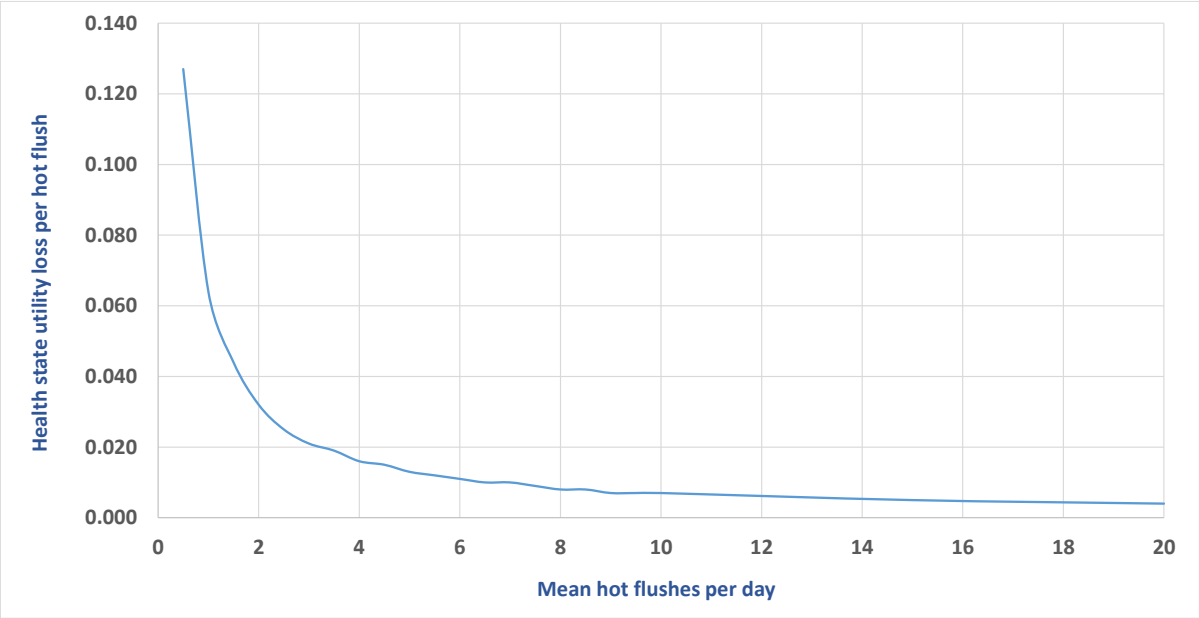
vi. Varying the health state utility loss from hot flushes

The absolute risks of breast cancer and VTE are small but the presence of vasomotor symptoms defines the population covered by the model. Therefore, the a priori expectation is that the health state utility loss assigned to hot flushes would be an important driver of cost-effectiveness. If the health state utility loss from vasomotor symptoms was trivial then even an intervention that greatly reduced their frequency would only have a limited QALY gain. Conversely, if vasomotor symptoms can produce a relatively large health state utility loss then the greater the potential for effective interventions to produce QALY gains at acceptable cost to the NHS.

As the previous analyses have suggested that non-oral oestrogen and progestogen can be considered a cost-effective short term treatment for menopausal symptoms, increasing the health state utility from flushing would only strengthen that conclusion. Therefore, in this analysis the health state utility loss for flushing was halved from its base case value and the threshold for mean flushes per day was established for non-oral oestrogen and progestogen to be considered cost-effective at this lower health state utility loss. It was found that when holding all other model inputs constant at their base case value that non-oral oestrogen and progestogen became more cost-effective than black cohosh when the mean flushes per day was six or more when assessed by the net mean benefit using a willingness to pay of £20,000 per QALY. The threshold for cost-effectiveness of non-oral oestrogen and progestogen in terms of mean hot flushes per day was the same when a willingness to pay of £30,000 per QALY was used.

Figure 13 shows the threshold health state utility according to symptom severity with health state utilities greater than or equal to the level indicated by the line being where non-oral and oestrogen is a cost-effective treatment for this level of symptom severity.

Figure 13: Threshold health state utility for cost-effectiveness at different symptom severity



vii. Varying the QALY loss from bleeding

If the duration of bleeding is reduced from 60 months to 24 months then this improves the cost-effectiveness of non-oral oestrogen and progestogen when compared to black cohosh, the next most effective treatment, with the ICER falling to £16,370 per QALY. Reducing the duration has the effect of reducing the QALY loss from bleeding, and given that HRT has higher rates of bleeding, this is an expected result.

However, if the QALY loss of bleeding was increased then the relative cost-effectiveness of non-oral oestrogen and progestogen would be expected to decline. In this analysis the health state utility loss was increased until non-oral oestrogen and progestogen was no longer more cost-effective than black cohosh in a woman experiencing a mean of 5 hot flushes per day. Providing the health state utility loss did not exceed 0.05 then non-oral oestrogen and progestogen remained cost-effective relative to black cohosh.

L.1.4.2 Women without a uterus

Base case analysis

This analysis suggests that non-oral oestradiol is the most cost-effective treatment of menopausal symptoms in women without a uterus. Not only does this have the highest net mean benefit and low ICER, it also has a relatively high probability of being the most cost-effective treatment. The results are shown in Table 23, Figure 14 and Figure 15. In this case some strategies are ruled out on the basis of extended dominance. This means that there are more effective and more costly strategies which have a lower ICER. The ICERs are calculated relative to the next best non-dominated (including extended dominance) strategy.

Table 23: Base case results for women without a uterus based on 10,000 simulations[a]

Treatment	Mean cost	Mean QALY	Mean Net Benefit	Probability cost-effective	ICER
No treatment	£0	0.0000	£0	13.6%	n/a
SSRIs/SNRIs	£57	0.0406	£754	6.6%	Extended dominance
Gabapentin	£61	0.0601	£1,142	10.3%	£1,007

Treatment	Mean cost	Mean QALY	Mean Net Benefit	Probability cost-effective	ICER
Oestradiol oral	£210	0.0897	£1,576	2.2%	Extended dominance
Isoflavones/Genistein/Soy	£314	0.1112	£1,911	1.5%	Extended dominance
Oestradiol non-oral	£357	0.1981	£3,606	39.1%	£2,149
Valerian Root	£438	0.0001	-£437	0.0%	Dominated
Black cohosh	£450	0.1674	£2,899	19.4%	Dominated
Multibotanicals	£486	0.0589	£692	3.5%	Dominated
Acupuncture	£545	0.1083	£1,621	3.9%	Dominated
Chinese herbal medicine	£2,033	-0.0019	-£2,072	0.0%	Dominated

(a) *Mean costs and QALYs calculated relative to no treatment*

Figure 14: Base case results for women without a uterus based on 10,000 simulations illustrated on the cost-effectiveness plane

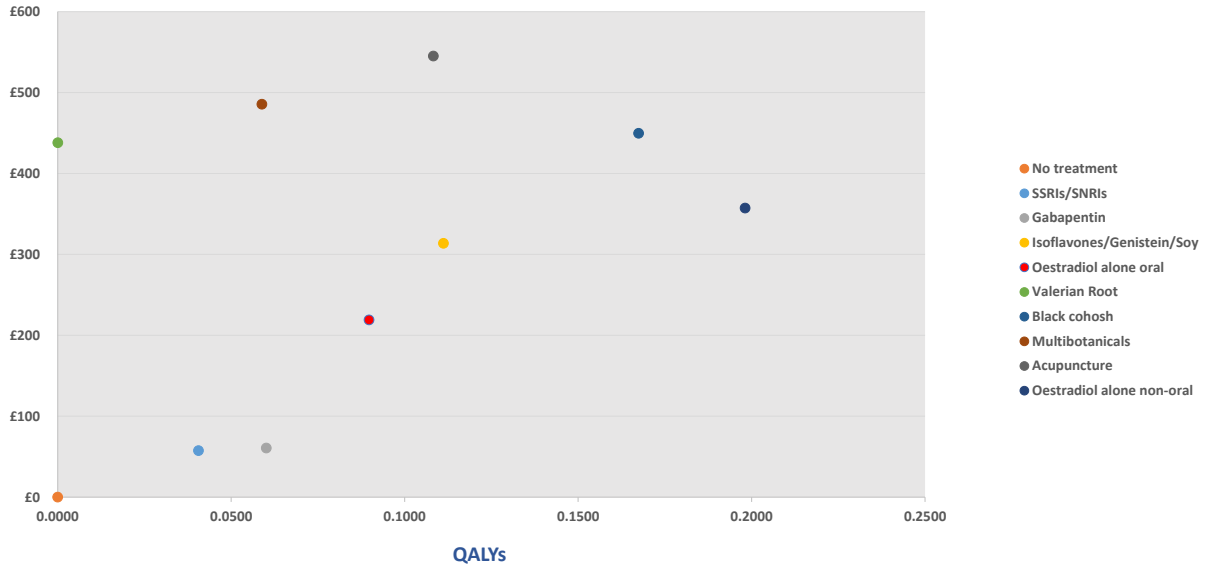


Figure 15: Cost-effectiveness plane showing results of 1,000 simulations for the base case analysis in women without a uterus

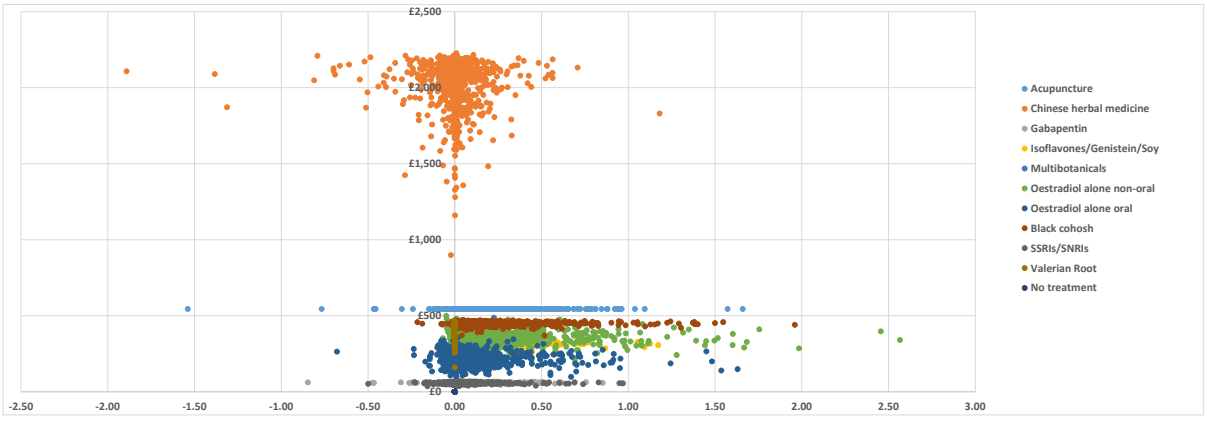
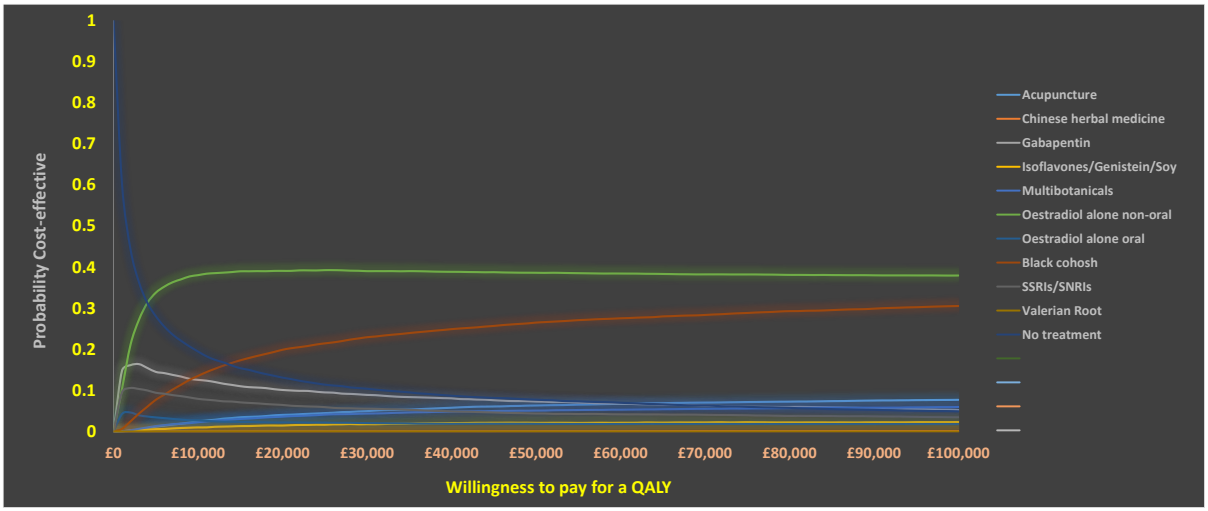


Figure 16 shows the CEAC for the model population without a uterus and with base inputs. Non-oral oestradiol has the highest probability of being the most cost-effective strategy unless the decision maker has a very low willingness to pay for a QALY. It also suggests that no treatment is always the cheapest alternative

Figure 16: Cost-effectiveness acceptability curve of base case analysis in women without a uterus based on 10,000 simulations

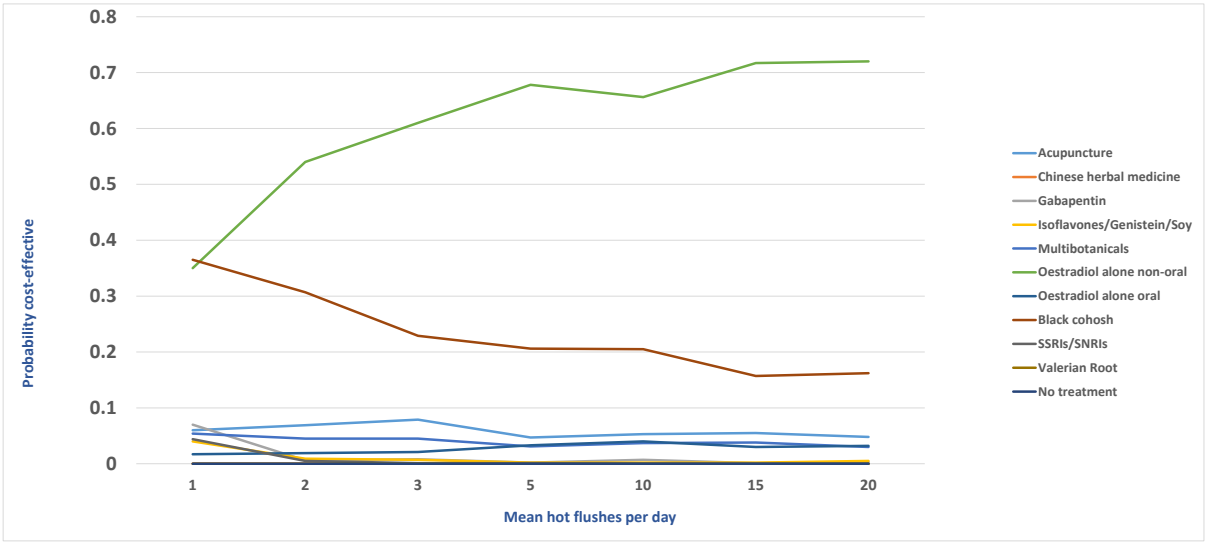


Sensitivity analysis

i. Disease severity

In this sensitivity analysis the mean number of hot flushes was varied from a mean of 1 hot flush per day to a mean of 20 hot flushes per day. Non-oral oestradiol becomes the most cost-effective at a mean of 2 hot flushes per day. Figure 17 shows how the probability of different treatments being cost-effective varies with severity of symptoms as measured by the number of mean hot flushes per day.

Figure 17: Probability of different treatments being cost-effective varying the mean hot flushes per day in women without a uterus



ii. Varying the costs and QALYs from breast cancer

If the base case analysis is re-run but using a breast cancer cost of £30,000 non-oral oestradiol goes from dominating black cohosh to having an ICER of £1,100 per QALY, still comfortably within a willingness to pay threshold of £20,000 per QALY. The probability of non-oral oestradiol being the most cost-effective falls from 39% to 34%, but still much higher than any other treatment alternative.

For a woman with a mean of 5 hot flushes per day non-oestradiol remains cost-effective even if the breast cancer cost is assumed to be £30,000 and if the health state utility loss from breast cancer is increased to 0.40 providing the duration of the health state utility loss is 21 years or less.

iii. Varying the health state utility loss from hot flushes

Holding all other inputs constant at their base case values, non-oral oestradiol remains the most cost-effective treatment in women without a uterus having a mean 3 hot flushes per day providing the health state utility loss from a hot flush is 0.009 or greater.

L.1.4.3 Women with breast cancer

Base case analysis

Relatively few treatments were included in this analysis reflecting the limited number of studies for which there is data for this population. Table 24 and Figure 18 suggest that St John's Wort is the most cost-effective treatment alternative in this population, although gabapentin also has a relatively high probability of being cost-effective.

Table 24: Base case results for women without a uterus based on 10,000 simulations[a]

Treatment	Mean cost	Mean QALY	Mean Net Benefit	Probability cost-effective	ICER
No Treatment	£0	0.0000	£0	9.3%	n/a
Gabapentin	£28	0.0598	£1,168	52.9%	£474
SSRIs/SNRIs	£33	-0.1662	-£3,358	2.8%	Dominated
Isoflavones/Genistein/Soy	£263	-0.0337	-£938	2.3%	Dominated
St John's Wort	£459	0.0919	£1,379	32.7%	£13,435

(a) Mean costs and QALYs calculated relative to no treatment

Figure 18: Base case results for women with breast cancer based on 10,000 simulations illustrated on the cost-effectiveness plane

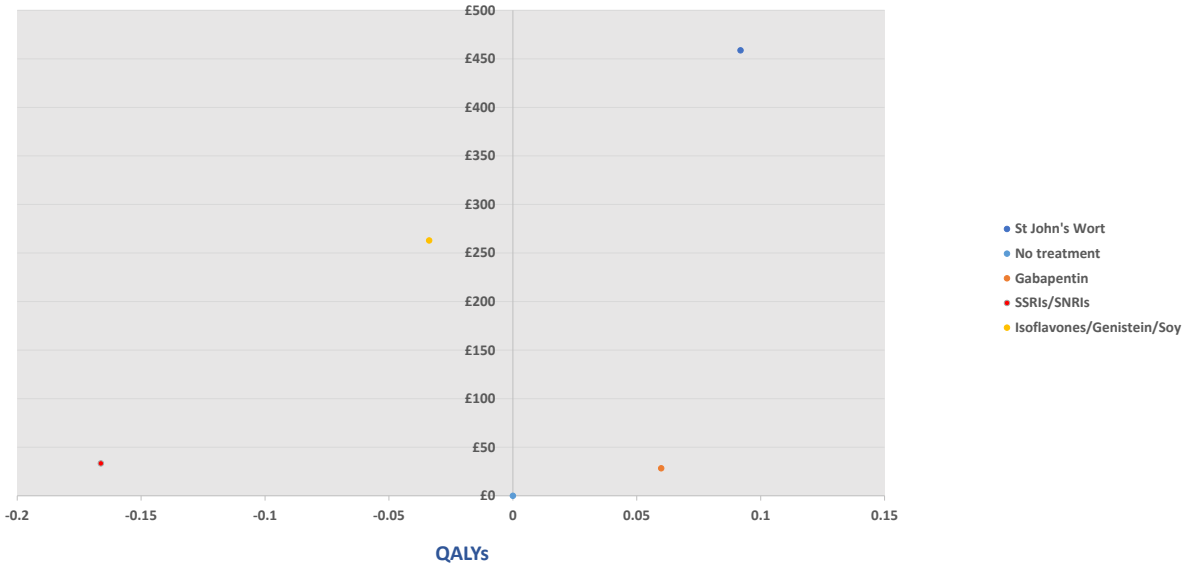


Figure 19: Cost-effectiveness plane showing results of 1,000 simulations for the base case analysis in women with breast cancer

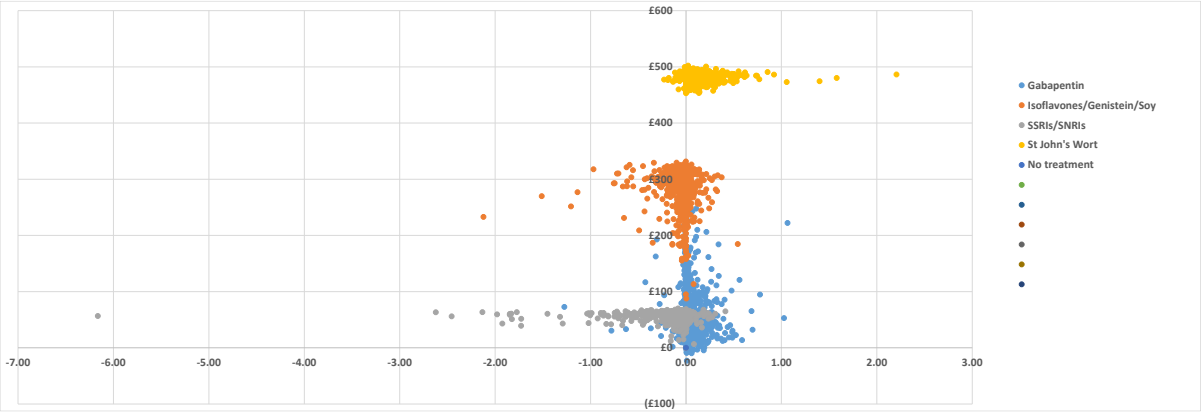
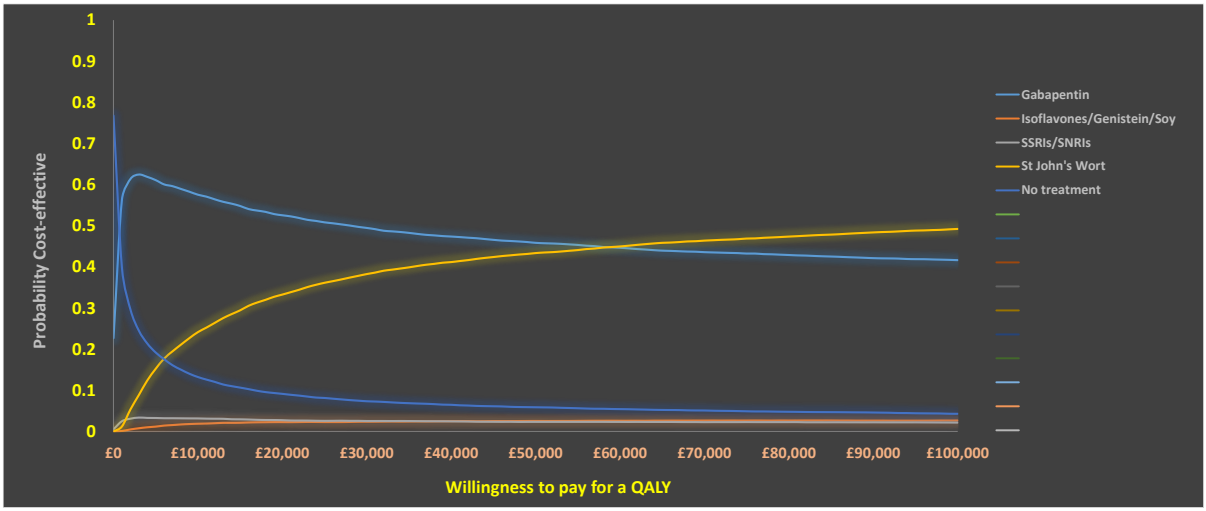


Figure 19 shows the CEAC for this analysis and suggests that gabapentin and no treatment have approximately the same probability of being the cheapest strategy. As the willingness to pay increases St John's Wort is increasingly favoured over Gabapentin.

Figure 20:Cost-effectiveness acceptability curve of base case analysis in women with breast cancer based on 10,000 simulations

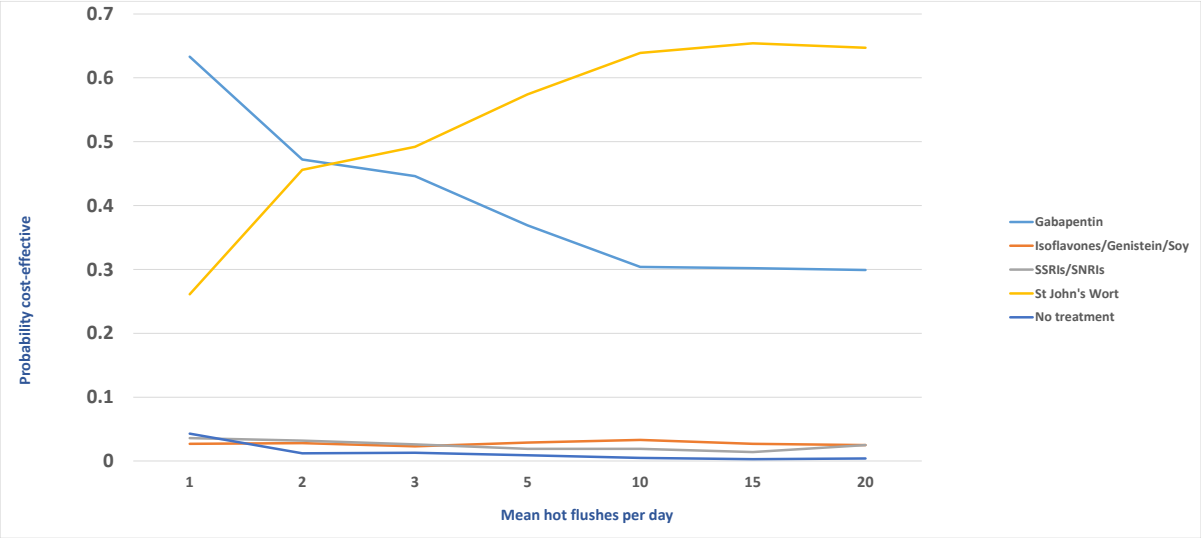


Sensitivity analysis

i. Disease severity

In this sensitivity analysis the mean number of hot flushes was varied from a mean of 1 hot flush per day to a mean of 20 hot flushes per day. This analysis suggested that St John's Wort was the most cost-effective treatment with 2 or more mean hot flushes per day in this population. Figure 21 shows how the probability of different treatments being cost-effective varies with severity of symptoms as measured by the number of mean hot flushes per day.

Figure 21: Probability of different treatments being cost-effective varying the mean hot flushes per day in women with breast cancer



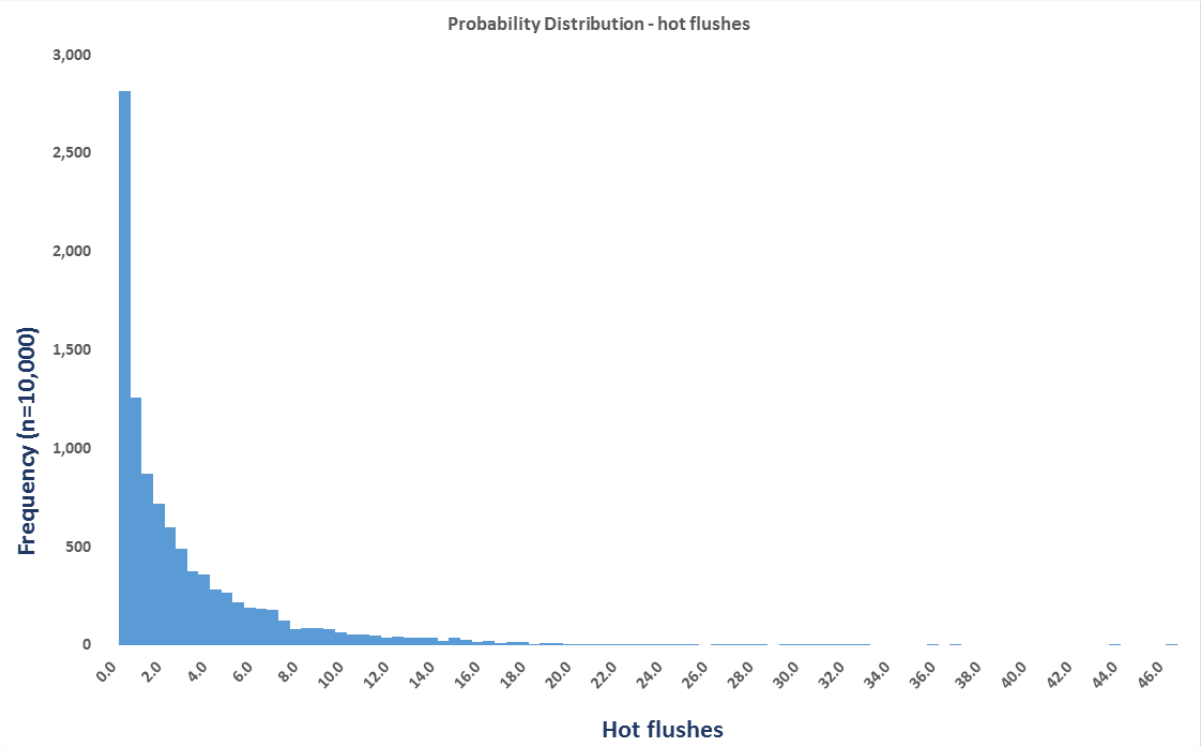
i. Varying health state utility loss from hot flushes

The base case analysis suggests that treatment compared to no treatment is cost-effective. This sensitivity analysis assessed the extent to which this finding was driven by the assumptions made with respect to the health state utility loss from hot flushes (as a proxy for vasomotor symptoms more generally). For a woman with a mean of 3 hot flushes per day and with breast cancer, gabapentin remained cost-effective providing the health state utility loss per hot flush was greater than 0.0005, a much lower value than used in the base case analysis.

L.1.5 Discussion

The results and sensitivity analysis presented in Appendix L.1.4 are organised by population with an unknown baseline of mean hot flushes but sampled from a distribution with a mean of 3 hot flushes per day (see Figure 22) or for a population where the mean hot flushes is known and can be varied.

Figure 22: Probability distribution for mean number of hot flushes per day used in the base case analysis



Using the population approach would perhaps be the most relevant for decision maker if the decision maker was unaware of symptom severity in the woman to be treated. The mean of 3 hot flushes per day is toward the more moderate end of the symptom severity spectrum and the lower symptom severity the less likely are women to seek treatment. Furthermore, the distribution is right skewed so that in a majority of the Monte Carlo simulations the number of mean hot flushes per day is less than three.

However, in practice the decision maker will often have a good idea of symptom severity as a result of taking a clinical history and therefore the deterministic approach may provide more meaningful results to decision makers, especially as cost-effectiveness of treatment, not surprisingly, improves with increasing symptom severity.

The models for the different populations (women with a uterus, women without a uterus and women with breast cancer) suggest that some treatment is cost-effective but it is important to be aware of the model limitations.

The model is underpinned by the network meta-analysis that was undertaken for this guideline. For each model a treatment was generally only included if it featured in at least one of the network meta-analyses undertaken for that population (see Appendix K). The one exception to this was for oral oestradiol in women without a uterus which the GDG thought needed to be included as this drug is usually the first line treatment in this population. Where network meta-analysis data was not available for a treatment outcome included in that model then it was necessary to extrapolate. That would sometimes mean assuming no relative treatment effect, assuming the same relative treatment effect as the same drug in a different population or assuming the same relative treatment effect as a different treatment in the same population. The GDG advised on these extrapolations choosing an alternative that they considered to be the best fit. Whilst, it would be preferable to have NMA data for all outcomes for each treatment, extrapolation is not necessarily an important limitation. For example, there is no biological reason to expect black cohosh to increase bleeding when compared to baseline. On the other hand some members of the GDG did express a view that

it was possible that there could be systematic differences in discontinuation and bleeding rates between oral and transdermal preparations of oestradiol and progestogen. In the model for women without a uterus the relative treatment effect of non-oral oestradiol was assumed to be the same as the relative treatment effect for non-oral oestradiol and progestogen in women with a uterus. Again, a member of the GDG commented that progestogen can enhance the effects of oestradiol on vasomotor symptoms and therefore this assumption may overestimate the relative treatment effect of non-oral oestradiol in this population.

It was not possible to include relatively new treatments such as bazedoxifene as there was no network meta-analysis data available for its effect on vasomotor symptoms.

It could be argued that the network meta-analysis poses a problem for inference. For example, it suggests that non-oral oestradiol and progestogen is significantly better than placebo for vasomotor symptoms using the conventional 5% level for statistical significance (see Appendix K) in women with a uterus. It also just fails to find that oral oestradiol and progestogen is significantly better than placebo. Does that mean that it can be inferred that non-oral oestradiol and progestogen is significantly better than oral oestradiol and progestogen? Unfortunately, that inference cannot be made although non-oral oestradiol and progestogen has a better point estimate of relative effect. As illustrated by this example, the rules of inference are not always helpful where decision makers have to make a choice between competing alternatives. For a decision maker seeking to maximise health outcomes it makes sense to choose the treatment which offers the greatest mean net mean benefit even when there may be a relatively high probability that it is not the most cost-effective treatment.

The model results presented give the probability of a treatment being considered cost-effective and this is a useful indicator of the uncertainty in outcomes. However, some care is sometimes needed in interpreting these results as it is not always the case that the treatment with the highest probability of being the most cost-effective treatment is also the most cost-effective as measured by the mean net benefit. The probability of a treatment being the most cost-effective is influenced by its relative treatment effect but also by the probability distribution surrounding that effect. For, example if one treatment in the analysis has much wider confidence intervals than another it may ‘win’ in those simulations sampled from its relatively long ‘left-hand’ tail. The greater uncertainty surrounding its true relative effect may result in a relatively high probability of it being cost-effective even when the expected mean net mean benefit is less, possibly considerably less, than other treatments. The corollary of these wide confidence intervals is that there is often a relatively high probability of it also being very poor value for money.

For women with a uterus, the model suggested that non-oral (which in practical terms means transdermal patches) oestradiol and progestogen was the most cost-effective treatment for women with menopausal symptoms for a frequency of 3 or more hot flushes per day. As the daily frequency of hot flushes increased so did the relative cost-effectiveness of non-oral oestradiol and progestogen and it also had an increasing probability of being the most cost-effective option. Black cohosh was slightly more cost-effective at a frequency of 2 hot flushes per day. These results were driven by the network meta-analysis on vasomotor symptoms in which black cohosh and non-oral oestradiol and progestogen had the best relative treatment effects and were the only two treatments significantly better than placebo at a 5% level of statistical significance. Figure 71 in Appendix K shows that black cohosh has the second lowest mean ratio for (relief of) vasomotor symptoms for women with a uterus and that this is a statistically significant result. Table 7 of Appendix K, shows that in the network meta-analysis, black cohosh has the second highest probability of being the most effective treatment and the probabilistic sensitivity analysis for the economic modelling, which gives the probability of black cohosh being the most cost-effective treatment, will reflect this. However, at lower symptom severity black cohosh has a higher probability of being cost-effective because of its lower cost relative to non-oral oestradiol and progestogen and lower risk of breast cancer and VTE. However, as symptom severity increases (where clinical

effectiveness becomes a more important driver of cost-effectiveness) the probability of black cohosh being the most cost-effective treatment declines as shown in Figure 9.

These findings were generally robust with respect to the assumptions made. Sensitivity analysis suggested that almost tripling the cost of breast cancer made little difference to the results. This is because, although a large cost for individuals, it represents a very small contribution to the total costs as the absolute risk are small. Increasing the utility losses associated with breast cancer also made little difference to results although a much higher QALY loss would make non-oral oestradiol start to look less cost-effective at relatively low levels of mean hot flushes per day.

Removing the VTE risk of tibolone and oral oestradiol and progestogen made a neglible impact on the model's results. As expected it marginally improved the cost-effectiveness of these two treatments but not sufficiently that they could be considered cost-effective relative to alternatives. This further illustrates the extent to which the model's results are driven primarily by the impact of treatment on vasomotor symptoms.

In women with a uterus the model results were only sensitive to treatment costs at relatively low levels of mean hot flushes per day. Furthermore, these cost parameters are often known and less subject to uncertainty than other model inputs. The model showed that at severe levels of mean hot flushes per day, that non-oral oestradiol and progestogen had a very high probability of being cost-effective even if black cohosh was assumed to cost only 20% of its base case value.

A simplifying assumption was made with respect to acupuncture that a treatment benefit would be maintained over the time horizon of the model after a single cycle (13 week) course of treatment. This was because it was difficult to estimate the intensity of treatment over a longer time frame. Of course, such an assumption is likely to lead to an underestimation of the treatment costs of acupuncture as it is likely that further treatment would be needed to maintain a treatment benefit. However, the base case model did not suggest that acupuncture was cost-effective even with this simplifying assumption and therefore the underlying cost-effectiveness conclusions of the model would not be affected if higher acupuncture costs were used.

One of the limitations of the model is that, beyond the first 2 cycles, discontinuation outcomes are not linked to other treatment outcomes. Perhaps perversely, increased discontinuation in the model after 2 cycles is linked with improved cost-effectiveness as it lowers cost but has no impact on QALYs. This is why the cost-effectiveness of non-oral oestradiol and progestogen improves relative to black cohosh with higher rates of end point discontinuation as the absolute difference in costs reduces because it is the more expensive treatment alternative. Less effect is observed in model outcomes when discontinuation is assumed to continue at the same rate as in the first two cycles because first, the final rate of discontinuation is much less than the 80% employed in one sensitivity analysis but also because a relative gap is maintained in discontinuation between black cohosh and non-oral oestradiol and progestogen.

The extent to which the lack of relationship between discontinuation and other model outcomes is an important limitation may depend to some extent on the reasons for discontinuation. If the women who discontinue after 6 months do so largely because symptoms have resolved or improved then it is less likely to be important. It should be noted that whilst the timeframe for this model is 5 years, the cost-effectiveness results are not likely to be different with a shorter timeframe, as costs and QALYs for most treatments in the analysis are essentially incurred at a virtually constant rate per cycle.

Given the importance of relative treatment effects on vasomotor symptoms in driving the cost-effectiveness results, it follows that the assumptions concerning the health state utility are likely to have an important bearing on treatment cost-effectiveness. However, the sensitivity analysis suggested that the health state utility loss from hot flushes would have to

be considerably lower than their base case values to alter cost-effectiveness conclusions even at relatively low levels of mean hot flushes per day. **Figure 13** also reinforces this as the threshold health state utility for non-oral oestrogen and progestogen to be cost-effective in women with a uterus would be considerably less than the base case health state utility loss from hot flushes beyond a mean of 4 hot flushes per day.

Similarly, the health state utility loss from bleeding would have to be considerably larger than the base case result in order to change model results.

In women without a uterus, non-oral oestradiol was markedly more cost-effective than any of the other treatment alternatives. However, there is an important caveat to this finding as the relative treatment effects were extrapolated from non-oral oestradiol and progestogen in women with a uterus which might cause the relative treatment effect to be over-estimated. The sensitivity analysis suggested that the model results would be substantially unchanged even with a number of less favourable input parameters.

In women with breast cancer, St John's Wort appeared as the most cost-effective treatment for women with a mean of 3 hot flushes per day or above. If the health state utility loss from hot flushes was massively over-estimated, gabapentin would still be considered cost-effective as a very cheap treatment and without incremental QALY losses due to breast cancer, bleeding or VTE.

It should be noted that this analysis only addressed first-line treatment and did not consider treatment switching as we did not have the data for the relative treatment effectiveness for 2nd line therapy.

L.1.6 Conclusion

The model suggests that transdermal oestradiol and progestogen was the most cost-effective treatment in women with a uterus and that is reflected in the recommendation of this guideline. However, the GDG noted that the network meta-analysis failed to reject a null hypothesis of no difference between transdermal and oral oestradiol and progestogen at the 5% significance level. Whilst, they believed that it was biologically plausible that transdermal patches would work better than oral preparations they didn't think the evidence was sufficiently strong to completely overturn current clinical practice and the use of much cheaper oral oestradiol and progestogen as the principle first line treatment. This is also reflected in the guideline recommendation which gives practitioners a choice between oral or transdermal preparations.

The model also suggested that non-oral oestradiol was cost-effective in women without a uterus although this model relied more heavily on extrapolated data. The guideline recommendations for women without a uterus mirror the recommendations for women with a uterus with a choice given between the use of oral and transdermal preparations with the same rationale.

Finally, the model suggested that St John's Wort could be considered as a cost-effective treatment for women with breast cancer and again this is reflected in a recommendation that advises women that it can be considered as a treatment option whilst highlighting the possibility of interaction with other medicines (e.g. tamoxifen).

Appendix M: Absolute risk references

M.1 Cardiovascular disease

Table 25: References for baseline risk and relative risks used to calculate absolute risk differences for CHD

		Reference for baseline risk over 7.5 years in the UK population: Weiner 2008				
		Past users	Current users	Treatment duration <5 years	Treatment duration 5–10 years	>5 years since stopping treatment
Women on oestrogen alone	RCT estimate	–	Manson 2013 (update)	–	–	Lacroix 2011 (the WHI reanalysis)
	Observational estimate	–	Grodstein 1996 (the NHS)	–	–	–
Women on oestrogen plus progestogen	RCT estimate	–	Manson 2013 (update)	–	–	Manson 2013 (update)
	Observational estimate	–	–	–	–	–
Women on any HRT	RCT estimate	–	Schierbeck 2012	–	–	Schierbeck 2012
	Observational estimate	Grodstein 2000 (the NHS)	Hedblad 2002; Lokkegaard 2008; Stram 2011; Grodstein 2000 (the NHS)			

HRT, hormone replacement therapy; RCT, randomised controlled trial

Table 26: References for baseline risk and relative risks used to calculate absolute risk differences for stroke

		Reference for baseline risk over 7.5 years in the UK population: Weiner 2008				
		Past users	Current users	Treatment duration <5 years	Treatment duration 5–10 years	>5 years since stopping treatment
Women on oestrogen alone	RCT estimate	–	Manson 2013 (update)	–	–	Manson 2013 (update)
	Observational estimate	–	Grodstein 1996 (the NHS)	–	–	–

		Reference for baseline risk over 7.5 years in the UK population: Weiner 2008				
Women on oestrogen plus progestogen	RCT estimate	–	Manson 2013 (update)	–	–	Manson 2013 (update)
	Observational estimate	–	Grodstein 2008 (the NHS)	–	–	–
Women on any HRT	RCT estimate	–	Schierbeck 2012	–	–	Schierbeck 2012
	Observational estimate	Grodstein 2000 (the NHS)	Grodstein 2000 (the NHS); Li 2006; Sourander 1998	–	Grodstein 2000 (the NHS)	–
HRT, hormone replacement therapy; RCT, randomised controlled trial						

M.2 Breast cancer

Table 27: References for baseline risk and relative risks used to calculate absolute risk differences for breast cancer

		Reference for baseline risk over 7.5 years in the UK population: [ONS data, 2010]				
		Past users	Current users	Treatment duration <5 years	Treatment duration 5–10 years	>5 years since stopping treatment
Women on oestrogen alone	RCT estimate	–	Manson 2013	–	–	Manson 2013
	Observational estimate	Willis 1996; Lund 2007; Saxena 2010; Sourander 1998	Bakken 2011; Lund 2007; Saxena 2010; Sourander 1998			
Women on oestrogen plus progestogen	RCT estimate		Beral, 2003; Bakken, 2011; Colditz, 1992; Willis, 1996; Fournier, 2008; Saxena, 2010; Bakken, 2004	Beral, 2003; Bakken, 2011; Colditz, 1992; Willis, 1996; Fournier, 2008	Willis 1996; Schairer 2000	
	Observational estimate					

		Reference for baseline risk over 7.5 years in the UK population: [ONS data, 2010]				
Women on any HRT	RCT estimate	–	Manson 2013	–	–	Lund 2007; Saxena 2010
	Observational estimate	Lund 2007; Saxena 2010	Bakken 2011; Lund 2007; Saxena 2010	Beral, 2003; Bakken, 2011; Fournier, 2008; Saxena, 2010; Schairer, 2000; Bakken, 2004	Beral, 2003; Bakken, 2011; Fournier, 2008	
HRT, hormone replacement therapy; RCT, randomised controlled trial						

M.3 Osteoporosis

Table 28: References for baseline risk and relative risks used to calculate absolute risk differences for osteoporosis

		Baseline risks calculated from control arm of each study				
		Past users	Current users	Treatment duration <5 years	Treatment duration 5–10 years	>5 years since stopping treatment
Women on any HRT	RCT estimate	–	Manson 2013	–	–	Lund 2007; Saxena 2010
	Observational estimate	Lund 2007; Saxena 2010	Bakken 2011; Lund 2007; Saxena 2010	Beral, 2003; Bakken, 2011; Fournier, 2008; Saxena, 2010; Schairer, 2000; Bakken, 2004	Beral, 2003; Bakken, 2011; Fournier, 2008	
HRT, hormone replacement therapy; RCT, randomised controlled trial						