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.

Copyright
National Collaborating Centre for Women's and Children's Health

Funding
The National Collaborating Centre for Women and Children’s Health was commissioned by the National Institute for Health and Care Excellence to undertake the work on this guideline.
# Contents

**Foreword** ........................................................................................................................................... 11

1 **Guideline summary** ......................................................................................................................... 14
   1.1 Guideline development group membership, NCC-WCH staff and acknowledgements .................. 14
   1.2 Care pathway ................................................................................................................................. 15
   1.3 Recommendations ......................................................................................................................... 18
   1.4 Research recommendations .......................................................................................................... 24
   1.5 Other versions of the guideline ..................................................................................................... 25
   1.6 Schedule for updating the guideline ............................................................................................ 25

2 **Development of the guideline** ........................................................................................................ 26
   2.1 What is a NICE clinical guideline? ................................................................................................. 26
   2.2 Remit ............................................................................................................................................... 26
   2.3 Who developed this guideline? ...................................................................................................... 27
   2.4 What this guideline covers ............................................................................................................ 27
      2.4.1 Groups that will be covered .................................................................................................... 27
      2.4.2 Key clinical issues that will be covered .................................................................................. 27
   2.5 What this guideline does not cover ............................................................................................... 28
      2.5.1 Groups that will not be covered ............................................................................................. 28
      2.5.2 Clinical issues that will not be covered .................................................................................. 28
   2.6 Relationships between the guideline and other NICE guidance .................................................. 29
      2.6.1 Related NICE guidance ........................................................................................................ 29

3 **Guideline development methodology** ............................................................................................. 30
   3.1 Developing the review questions and protocols .............................................................................. 30
   3.2 Searching for evidence ................................................................................................................... 33
      3.2.1 Clinical literature search ....................................................................................................... 33
   3.3 Reviewing and synthesising the evidence ...................................................................................... 34
      3.3.1 Methods of combining clinical studies .................................................................................... 34
      3.3.2 Type of studies ....................................................................................................................... 36
      3.3.3 Appraising the quality of evidence by outcomes ..................................................................... 37
      3.3.4 Use of absolute effect in decision making ............................................................................ 42
      3.3.5 Evidence statements .............................................................................................................. 43
      3.3.6 Evidence of cost effectiveness .............................................................................................. 43
   3.4 Developing recommendations ....................................................................................................... 44
      3.4.1 Research recommendations ..................................................................................................... 45
      3.4.2 Validation process ................................................................................................................... 45
      3.4.3 Updating the guideline .......................................................................................................... 45
      3.4.4 Disclaimer .............................................................................................................................. 45
      3.4.5 Funding ................................................................................................................................... 46
4 Diagnosis of perimenopause and menopause .......................................................... 47
  4.1 Review question .................................................................................................. 47
  4.2 Introduction to topic ......................................................................................... 47
  4.3 Clinical introduction ......................................................................................... 47
  4.4 Description of included studies ....................................................................... 48
    4.4.1 Studies looking at the diagnostic accuracy of age and menopausal symptoms ........................................................................................................ 48
    4.4.2 Studies looking at the diagnostic accuracy of biochemical measures ...... 48
    4.4.3 Studies looking at the diagnostic accuracy of ultrasound features ........ 49
    4.4.4 Studies looking at combination tools ....................................................... 49
  4.5 Clinical evidence profile .................................................................................. 49
  4.6 Economic evidence ......................................................................................... 53
  4.7 Evidence statements ....................................................................................... 53
  4.8 Evidence to recommendations ....................................................................... 56
    4.8.1 Relative value placed on the outcomes considered .................................. 56
    4.8.2 Consideration of clinical benefits and harms .......................................... 56
    4.8.3 Consideration of economic benefits and harms ....................................... 57
    4.8.4 Quality of evidence .................................................................................. 57
    4.8.5 Other considerations ................................................................................ 57
    4.8.6 Key conclusions ..................................................................................... 57
  4.9 Recommendations .......................................................................................... 58

5 Classification systems for the diagnosis of menopause ........................................ 59
  5.1 Review question .............................................................................................. 59
  5.2 Introduction to topic ....................................................................................... 59
  5.3 Clinical introduction ....................................................................................... 59
  5.4 Description of included studies ...................................................................... 59
  5.5 Clinical evidence profile ................................................................................. 59
  5.6 Economic evidence ....................................................................................... 59
  5.7 Evidence statements ...................................................................................... 59
  5.8 Evidence to recommendations ....................................................................... 60
    5.8.1 Relative value placed on the outcomes considered .................................. 60
    5.8.2 Consideration of clinical benefits and harms .......................................... 60
    5.8.3 Consideration of economic benefits and harms ....................................... 60
    5.8.4 Quality of evidence .................................................................................. 60
    5.8.5 Other considerations ................................................................................ 60
    5.8.6 Key conclusions ..................................................................................... 60

6 Information and advice ....................................................................................... 61
  6.1 Review question .............................................................................................. 61
  6.2 Introduction to topic ....................................................................................... 61
  6.3 Clinical Introduction ....................................................................................... 61
6.4 Description of included studies ................................................................. 62
6.5 Clinical evidence profile ............................................................................. 63
6.6 Economic evidence ..................................................................................... 63
6.7 Evidence statements .................................................................................. 64
6.8 Evidence to recommendations ................................................................... 65
  6.8.1 Relative value placed on the outcomes considered ......................... 65
  6.8.2 Consideration of clinical benefits and harms .................................. 65
  6.8.3 Consideration of economic benefits and harms .............................. 66
  6.8.4 Quality of evidence ............................................................................. 67
  6.8.5 Other considerations .......................................................................... 67
  6.8.6 Key conclusions .................................................................................. 67
6.9 Recommendations ...................................................................................... 67

7 Managing short-term symptoms .................................................................. 69
  7.1 Review question ....................................................................................... 69
  7.2 Introduction to topic ................................................................................ 69
  7.3 Clinical introduction ................................................................................. 70
  7.4 Description of included studies ............................................................... 72
  7.5 Clinical Evidence profile .......................................................................... 88
  7.6 Economic evidence .................................................................................. 88
  7.7 Evidence statements ................................................................................. 92
    7.7.1 Evidence summary from the NMA ............................................... 92
    7.7.2 Evidence summary from the pair-wise comparisons ..................... 93
  7.8 Evidence to recommendations ................................................................. 99
    7.8.1 Relative value placed on the outcomes considered ..................... 99
    7.8.2 Consideration of clinical benefits and harms .............................. 99
    7.8.3 Consideration of health benefits and resource uses .................... 102
    7.8.4 Quality of evidence ......................................................................... 103
    7.8.5 Other considerations ...................................................................... 103
    7.8.6 Key conclusions ............................................................................. 103
  7.9 Recommendations .................................................................................... 104
  7.10 Research recommendations ..................................................................... 105
  7.11 Urogenital atrophy .................................................................................. 107
    7.11.1 Review question .............................................................................. 107
    7.11.2 Introduction to topic ...................................................................... 107
    7.11.3 Clinical introduction ...................................................................... 107
    7.11.4 Description of included studies ................................................... 108
    7.11.5 Clinical evidence profile ............................................................... 108
    7.11.6 Evidence statements ...................................................................... 109
    7.11.7 Health economics profile ............................................................. 111
    7.11.8 Evidence to recommendations ..................................................... 112
## Contents

### 8  Review and referral ..................................................................... 115

- 8.1 Review question ........................................................................ 115
- 8.2 Introduction to topic .................................................................. 115
- 8.3 Clinical Introduction .................................................................. 115
- 8.4 Description of included studies .................................................. 115
- 8.5 Evidence statements ................................................................... 115
- 8.6 Evidence to recommendations ...................................................... 116
  - 8.6.1 Relative value placed on the outcomes considered .................. 116
  - 8.6.2 Consideration of clinical benefits and harms ............................. 116
  - 8.6.3 Consideration of economic benefits and harms ......................... 116
  - 8.6.4 Quality of evidence ................................................................ 116
  - 8.6.5 Other considerations ................................................................ 116
  - 8.6.6 Key conclusions ..................................................................... 117
- 8.7 Recommendations ....................................................................... 117

### 9  Starting and stopping HRT ............................................................... 118

- 9.1 Review question ........................................................................ 118
- 9.2 Introduction to topic ................................................................... 118
- 9.3 Clinical introduction ................................................................... 118
- 9.4 Description of included studies ..................................................... 118
- 9.5 Clinical evidence profile .............................................................. 119
- 9.6 Economic evidence ...................................................................... 119
- 9.7 Evidence statements ................................................................... 120
- 9.8 Evidence to recommendations ...................................................... 120
  - 9.8.1 Relative value placed on the outcomes considered .................. 120
  - 9.8.2 Consideration of clinical benefits and harms ............................. 121
  - 9.8.3 Consideration of economic benefits and harms ......................... 121
  - 9.8.4 Quality of evidence ................................................................ 121
  - 9.8.5 Other considerations ................................................................ 122
  - 9.8.6 Key conclusions ..................................................................... 122
- 9.9 Recommendations ....................................................................... 122

### 10  Long-term benefits and risks of HRT .............................................. 123

- 10.1 Venous thromboembolism ........................................................... 123
  - 10.1.1 Review question .................................................................... 123
  - 10.1.2 Introduction to topic ............................................................... 123
  - 10.1.3 Clinical introduction ............................................................... 123
  - 10.1.4 Description of included studies ................................................ 124
  - 10.1.5 Clinical evidence profile ........................................................ 124
  - 10.1.6 Economic evidence ............................................................... 124
  - 10.1.7 Evidence statements ............................................................. 125
# Contents

10.1.8 Evidence to recommendations .......................................................... 127  
10.1.9 Recommendations .......................................................................... 129  
10.1.10 Research Recommendations .......................................................... 130  

10.2 Cardiovascular disease ........................................................................ 131  
10.2.1 Review question ................................................................................ 131  
10.2.2 Introduction to topic .......................................................................... 131  
10.2.3 Clinical introduction .......................................................................... 131  
10.2.4 Description of included studies ......................................................... 131  
10.2.5 Clinical evidence profile .................................................................... 134  
10.2.6 Economic evidence ........................................................................... 134  
10.2.7 Evidence statements .......................................................................... 135  
10.2.8 Evidence to recommendations .......................................................... 139  
10.2.9 Recommendations ............................................................................ 141  

10.3 Development of Type 2 Diabetes .......................................................... 143  
10.3.1 Review question ................................................................................ 143  
10.3.2 Introduction to topic .......................................................................... 143  
10.3.3 Clinical introduction .......................................................................... 143  
10.3.4 Description of included studies ......................................................... 143  
10.3.5 Clinical evidence profile .................................................................... 145  
10.3.6 Economic evidence ........................................................................... 145  
10.3.7 Evidence statements .......................................................................... 145  
10.3.8 Evidence to recommendations .......................................................... 146  
10.3.9 Recommendations ............................................................................ 147  

10.4 T2D management – control of blood sugar ........................................... 147  
10.4.1 Review question ................................................................................ 147  
10.4.2 Introduction to topic .......................................................................... 147  
10.4.3 Clinical introduction .......................................................................... 148  
10.4.4 Description of included studies ......................................................... 148  
10.4.5 Clinical evidence profile .................................................................... 150  
10.4.6 Economic evidence ........................................................................... 150  
10.4.7 Evidence statements .......................................................................... 150  
10.4.8 Evidence to recommendations .......................................................... 150  
10.4.9 Recommendations ............................................................................ 152  

10.5 Breast Cancer ....................................................................................... 152  
10.5.1 Review question ................................................................................ 152  
10.5.2 Introduction to topic .......................................................................... 152  
10.5.3 Clinical Introduction .......................................................................... 153  
10.5.4 Description of included studies ......................................................... 153  
10.5.5 Clinical evidence profile .................................................................... 155  
10.5.6 Evidence statements .......................................................................... 156
## Contents

### 10.5 Menopause

10.5.7 Health economics profile ................................................................. 157  
10.5.8 Evidence to recommendations ......................................................... 158  
10.5.9 Recommendations ............................................................................. 160  
10.5.10 Research Recommendations ......................................................... 161

### 10.6 Osteoporosis ................................................................. 162  
10.6.1 Review question ................................................................................. 162  
10.6.2 Introduction to topic ........................................................................... 163  
10.6.3 Clinical introduction .......................................................................... 163  
10.6.4 Description of included studies ......................................................... 164  
10.6.5 Clinical evidence profile ..................................................................... 175  
10.6.6 Economic evidence ............................................................................ 175  
10.6.7 Evidence statements .......................................................................... 175  
10.6.8 Evidence to recommendations ......................................................... 177  
10.6.9 Recommendations ............................................................................. 178

### 10.7 Dementia ................................................................. 179  
10.7.1 Review question ................................................................................. 179  
10.7.2 Introduction to topic .......................................................................... 179  
10.7.3 Clinical introduction .......................................................................... 180  
10.7.4 Description of included studies ......................................................... 180  
10.7.5 Summary of included studies ............................................................ 181  
10.7.6 Clinical evidence profile ..................................................................... 182  
10.7.7 Economic evidence ............................................................................ 182  
10.7.8 Evidence statements .......................................................................... 182  
10.7.9 Evidence to recommendations ......................................................... 183  
10.7.10 Recommendations .......................................................................... 184  
10.7.11 Research Recommendations ........................................................ 184

### 10.8 Loss of muscle mass (sarcopenia) ......................................................... 185  
10.8.1 Review question ................................................................................. 185  
10.8.2 Introduction to topic .......................................................................... 185  
10.8.3 Clinical introduction .......................................................................... 186  
10.8.4 Description of included studies ......................................................... 186  
10.8.5 Clinical evidence profile ..................................................................... 186  
10.8.6 Economic evidence ............................................................................ 186  
10.8.7 Evidence statements .......................................................................... 187  
10.8.8 Evidence to recommendations ......................................................... 187  
10.8.9 Recommendations ............................................................................. 188

### 11 Premature ovarian insufficiency ................................................................. 189

#### 11.1 Diagnosis of premature ovarian insufficiency ........................................ 189  
11.1.1 Review question ................................................................................. 189  
11.1.2 Clinical introduction .......................................................................... 189
11.1 Description of included studies ................................................................. 190
11.1.3 Description of included studies ................................................................. 190
11.1.4 Clinical evidence profile ........................................................................... 190
11.1.5 Economic evidence .................................................................................. 192
11.1.6 Evidence statements ................................................................................ 193
11.1.7 Evidence to recommendations ................................................................. 194
11.1.8 Recommendations ................................................................................... 195
11.1.3 Description of included studies ................................................................. 190
11.1.4 Clinical evidence profile ........................................................................... 190
11.1.5 Economic evidence .................................................................................. 192
11.1.6 Evidence statements ................................................................................ 193
11.1.7 Evidence to recommendations ................................................................. 194
11.1.8 Recommendations ................................................................................... 195
11.2 Management of premature ovarian insufficiency .......................................... 196
11.2.1 Introduction to topic ................................................................................ 196
11.2.2 Review question ....................................................................................... 196
11.2.3 Clinical introduction ............................................................................... 196
11.2.4 Description of included studies ................................................................ 196
11.2.5 Clinical evidence profile ......................................................................... 198
11.2.6 Economic evidence ............................................................................... 198
11.2.7 Evidence statements ............................................................................... 199
11.2.8 Evidence to recommendations ............................................................... 199
11.2.9 Recommendations .................................................................................. 200
11.2.10 Research recommendations ................................................................ 201
12 References .................................................................................................... 203
13 Glossary ......................................................................................................... 256
14 Terms, acronyms and abbreviations ................................................................. 270
15 Appendices .................................................................................................... 273
1 Foreword

2 There were more than 11 million women over the age of 45 years in the UK according to the Office of National Statistics 2011 census. This number has been steadily increasing and is forecast to continue to rise. The associated increase in the number of women going through the menopause is expected to result in more new referrals to secondary care – both of women needing short-term symptom control and of women who have associated long-term health issues.

3 Menopause is a biological stage in a woman’s life when she is no longer fertile and is marked by the cessation of menstruation. A woman is defined as postmenopausal from 1 year after her last period. The changes associated with menopause and the perimenopause (the years leading up to it), occur when ovarian function diminishes and ceases. This includes both the cessation of egg (oocyte) maturation and of sex hormone (principally oestrogen and progesterone) secretion.

4 Unlike in the male where sperm production continues into old age, women have a finite number of oocytes at birth, which decline with each menstrual cycle. The menopause is characterised by the eventual depletion of the oocyte store and cessation of menstruation. Menstrual cycle irregularity often occurs before periods stop completely.

5 Most tissues contain oestrogen receptors through which the hormone exerts its effects. The most immediate changes resulting from reduced oestrogen levels are evident in the regulation of the menstrual cycle. However, oestrogen depletion associated with the menopause has many other effects on the body – for example, causing vasomotor, musculoskeletal, urogenital and psychological symptoms. It has also been shown to have an impact on the function of other systems in later life, including bone and the cardiovascular system. Oestrogen depletion explains some of the differences in the incidence of osteoporosis between men and women.

6 Perimenopause, also called the menopausal transition or climacteric, is the interval in which a woman has irregular cycles of ovulation and menstruation before the menopause. Within the UK population, the mean age of the natural menopause is 51 years, although this can vary between different ethnic groups.

7 Premature ovarian insufficiency (also known as premature ovarian failure or premature menopause) is usually defined as menopause occurring before the age of 40 years. It can occur naturally or iatrogenically (that is, as a result of treatment). Premature ovarian insufficiency (POI) and early perimenopause (menopause between the ages of 40 and 45 years) are associated with an increased risk of mortality, and with serious morbidity including cardiovascular disease (CVD), neurological disease, psychiatric disorders and osteoporosis.

8 Lower socioeconomic status has been associated with POI.

9 Many women experience a range of symptoms during the menopause and perimenopause and these symptoms are often short lived and lessen or disappear over time. The most common include vasomotor symptoms (for example hot flushes and sweats), effects on mood (for example low mood) and urogenital symptoms (for example vaginal dryness). Of women responding to a postal survey carried out in Scotland in 2009 about symptoms experienced in the previous month, 47% reported hot flushes, 46% reported night sweats and 26% reported vaginal dryness. The USA Study of Women’s Health Across the Nation reported in 2009 that, on average, African–American women had more hot flushes than white women, and Asian women (Japanese or Chinese) had the fewest hot flushes of all ethnic groups surveyed. The same study reported that early menopause (between 40 and 45 years of age) affected 3.7% of African–American women, 2.9% of white women, 2.2% of Chinese women and 0.8% of Japanese women.
Postmenopausal women are at increased risk of a number of long-term conditions, such as osteoporosis, CVD and changes in the vagina and bladder. These occur because of natural aging as well as oestrogen depletion.

During the latter part of the last century, hormone replacement therapy (HRT), also known as hormone therapy (HT) and menopausal hormone therapy (MHT) was advocated for both symptom relief and chronic disease prevention. This followed publication of several observational studies suggesting a decrease in the incidence of CVD, osteoporosis and dementia among other conditions of age. However, 2 landmark studies, the Women's Health Initiative (2002) and the Million Women Study (2003), reported on the risks and benefits associated with the use of hormone replacement therapy (HRT). The publication of these 2 studies was associated with a significant reduction in women's use of HRT in the UK. A retrospective GP database study (2010) reported that 18% of women aged 45–64 years consulted their GP at least once in 1996 for menopause-related symptoms, but this fell to 10% of women in 2005. Furthermore, a cross-sectional study in 2012 found that more than 60% of women managed their menopausal symptoms without any contact with healthcare professionals, often through social support and obtaining advice from friends, family and the internet.

Variations in consultation patterns for menopausal symptoms depend on many factors, including cultural, ethnic, educational and psychosocial factors, as well as the impact of the symptoms on the women. However, it is currently thought that more than one-third of all women want more support for managing menopausal symptoms from their GP or practice nurse.

The information and support offered to women during and after the menopause is thought to be variable and, for many, inadequate. A UK-based survey published in 2007 indicated that most women would welcome more information about the menopause. To improve the information provided, and to facilitate women being able to make an informed choice, some professional groups have suggested that all women should be invited for a health and lifestyle consultation when they reach the age of 50 years, which would include a discussion of menopausal symptoms and possible long-term sequelae of oestrogen depletion.

Treatments that have been used for menopause-related symptoms include lifestyle advice, HRT, herbal remedies, other complementary (alternative) therapies and antidepressants. In an internet survey (hosted at www.menopausematters.org.uk between 2005 and 2006), nearly three-quarters of women reported they did not know enough about HRT to make informed choices, 85% felt they did not know enough to make informed choices about alternative therapies for menopause-related symptoms, and 95% said they would try alternative therapies before HRT in the belief that they are more ‘natural’ and because of concern over the health risks of HRT. The use of HRT in the UK is strongly linked to socioeconomic status, with women of lower socioeconomic status being less likely to use HRT. Inequalities in referral rates have been associated with geography and age. There is also published evidence that physician speciality is significantly associated with HRT use. For example, in the USA women receiving care from gynaecologists are 2.6 times more likely to use HRT than women receiving care from family physicians.

There is no consensus about the long-term benefits and risks of HRT. The Women’s Health Initiative initially reported that HRT increased the risk of having a cardiovascular event as well as the incidence of breast cancer although it prevented osteoporotic fractures and colon cancer. However, the association between HRT and CVD has since been disputed and the results are now being interpreted differently.

In summary, a large number of women in the UK experience menopausal symptoms, which in many cases can significantly affect their quality of life. It is probable that a minority of these women seek medical treatment, and for those who do there is considerable variation in the help available with many being told that the symptoms will get better with time. Since symptoms may often continue for 7 years or more, this advice is inappropriate and help
should be offered where possible. Women need to know about the available options, their
risks and benefits and be empowered to become part of the decision-making process. The
need for the Guideline was recognised by the Department of Health and aims to provide
advice for both health care professionals and women regarding the menopause and the way
symptom relief can be achieved. It not only covers women who go through the menopause in
middle age, but also those with POI and for whom hormones are not appropriate including
women with, or at high risk of breast cancer. It covers the diagnosis and optimal clinical
management of menopause-related symptoms, including hormonal and non-hormonal
therapies. Attention is also given to the contentious issue of the impact of HRT on chronic
disease prevention although other, established treatments for CVD and osteoporosis in
particular, are not covered.
1 Guideline summary

1.1 Guideline development group membership, NCC-WCH staff and acknowledgements

Table 1: Guideline development group members

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary Ann Lumsden</td>
<td>Professor of Medical Education and Gynaecology Reproductive and Maternal Medicine Head of University of Glasgow Campus, Glasgow Royal Infirmary</td>
</tr>
<tr>
<td>Melanie Davies (until November 2014)</td>
<td>Consultant Obstetrician and Gynaecologist, University College London Hospitals</td>
</tr>
<tr>
<td>Nicholas Panay</td>
<td>Consultant Gynaecologist, Specialist in Reproductive medicine, Queen Charlotte's and Chelsea &amp; Chelsea and Westminster Hospitals, London</td>
</tr>
<tr>
<td>Imogen Shaw</td>
<td>GP, Finchingfield, Essex</td>
</tr>
<tr>
<td>Sally Hope</td>
<td>GP, Oxford, Oxfordshire</td>
</tr>
<tr>
<td>Anthony Parsons</td>
<td>Consultant Community Gynaecologist, Coventry and Warwickshire Partnership Trust</td>
</tr>
<tr>
<td>Deborah Holloway</td>
<td>Nurse Consultant Gynaecology, Guys and St Thomas’s NHS Foundation Trust</td>
</tr>
<tr>
<td>Prunella Neale</td>
<td>Practice Nurse, Herschel Medical Centre, Slough</td>
</tr>
<tr>
<td>Terry Aspray</td>
<td>Consultant Physician, Musculoskeletal Unit, Freeman Hospital</td>
</tr>
<tr>
<td>Sara Moger</td>
<td>Lay member</td>
</tr>
<tr>
<td>Claire Bowring</td>
<td>Lay member</td>
</tr>
<tr>
<td>Deborah Keatley</td>
<td>Lay member</td>
</tr>
<tr>
<td>Christine West (from Jan 2015)</td>
<td>Consultant Gynaecologist, Simpson Centre for Reproductive Health, Royal Infirmary of Edinburgh.</td>
</tr>
</tbody>
</table>

Expert advisers

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rebecca Hardy</td>
<td>Programme Leader for MRC Unit for Lifelong Health and Aging, University College London</td>
</tr>
<tr>
<td>Peter Collins</td>
<td>Professor of Clinical Cardiology, Imperial College London</td>
</tr>
<tr>
<td>Myra Hunter</td>
<td>Professional Lead for Clinical Health Psychology, South London and Maudsley Foundation Trust</td>
</tr>
<tr>
<td>Charlotte Coles</td>
<td>Consultant Clinical Oncologist, Addenbrooke’s Hospital, Cambridge</td>
</tr>
<tr>
<td>Adrian Harnett</td>
<td>Consultant Clinical Oncologist, Norfolk and Norwich University Hospital</td>
</tr>
</tbody>
</table>

Table 2: NCC-WCH staff

<table>
<thead>
<tr>
<th>Name</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammati Sarri (from October 2014)</td>
<td>Guideline Lead and Senior Research Fellow</td>
</tr>
<tr>
<td>Melanie Davies (from December 2014)</td>
<td>Clinical Director</td>
</tr>
<tr>
<td>Annabel Flint (from June 2014)</td>
<td>Project Manager</td>
</tr>
<tr>
<td>Yelan Guo (from March 2014)</td>
<td>Research Fellow</td>
</tr>
<tr>
<td>Sadia Janjua (from July 2014)</td>
<td>Research Fellow</td>
</tr>
<tr>
<td>Amy Wang (from June 2014)</td>
<td>Research Fellow</td>
</tr>
<tr>
<td>Hugo Pedder (from September 2014)</td>
<td>Statistician</td>
</tr>
<tr>
<td>Paul Jacklin (from January 2015)</td>
<td>Senior Health Economist</td>
</tr>
<tr>
<td>Name</td>
<td>Role</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>Omnia Abdulrazeg (from September to December 2014)</td>
<td>Research Fellow</td>
</tr>
<tr>
<td>Zosia Beckles (from November 2014)</td>
<td>Information Scientist</td>
</tr>
<tr>
<td>Rosalind Lai (until October 2014)</td>
<td>Information Scientist</td>
</tr>
<tr>
<td>David James (until November 2014)</td>
<td>Clinical Director</td>
</tr>
<tr>
<td>Hannah Rose Douglas (until May 2014)</td>
<td>Senior Health Economist and Guideline Lead</td>
</tr>
<tr>
<td>David Bevan (until January 2014)</td>
<td>Project Manager</td>
</tr>
<tr>
<td>Hugh McGuire (until March 2014)</td>
<td>Senior Research Fellow</td>
</tr>
<tr>
<td>Katie Webster</td>
<td>Research Fellow</td>
</tr>
<tr>
<td>Rupert Franklin (until June 2014)</td>
<td>Project Manager</td>
</tr>
<tr>
<td>Jiri Chard (until August 2014)</td>
<td>Guideline Lead</td>
</tr>
<tr>
<td>Fiona Caldwell (from January until July 2014)</td>
<td>Research Assistant</td>
</tr>
<tr>
<td>Sabina Sanghera (from April until August 2014)</td>
<td>Health Economist</td>
</tr>
<tr>
<td>Paul Mitchell (August 2014)</td>
<td>Health Economist</td>
</tr>
<tr>
<td>Setor Kunutsor (from May until November 2014)</td>
<td>Research Fellow</td>
</tr>
<tr>
<td>Katherine Cullen (from October 2014 until January 2015)</td>
<td>Health Economist</td>
</tr>
</tbody>
</table>

1 Additional support was received from Taryn Krause, Timothy Reeves, Kathryn Coles, Nitara Prasannan, and Sarah Bailey.

2 Ebenezer Ademisoye, Nitara Prasannan, and Sarah Bailey.

1.23 Care pathway
Figure 1: Care pathway
Woman presents with menopausal symptoms

- **Diagnosis of menopause**
  - Yes
  - No

  - **Assess risk factors for treatment**
    - Yes
    - No

  - **Discuss benefits and risks of treatments**
    - Yes
    - No

  - **Start treatment?**
    - Yes
    - No

  - **Interim monitoring at 3 months**
    - Yes
    - No

    - **Resolution of symptoms**
      - Yes
      - No

    - **Annual Review**
      - Yes
      - No

    - **Reassess risk factors (including age)**
      - Yes
      - No

    - **Suitable to continue treatment**
      - Yes
      - No

    - **Stop treatment (assess continuing health needs)**

- **Consider differential diagnosis**
  - Yes
  - No

- **Within normal range**

- **Investigations FSH, LH, oestradiol x 2**

- **Abnormal**

  - **Consider diagnosis of POI**
    - Yes
    - No

  - **Consider referral to specialist service**

  - **Refer to specialist service**

- **Resolution of symptoms**

- **Treatment failure**

- **Consider alternatives**

- **Consider referral to specialist service**

- **Annual Review**

- **Reassess risk factors (including age)**

- **Suitable to continue treatment**

- **Stop treatment (assess continuing health needs)**

- **Consider referral to specialist service**

- **Inform provision and health promotion**
1.3 Recommendations

1. Diagnose the following without laboratory tests in otherwise healthy women aged over 45 years with menopausal symptoms:
   - perimenopause based on vasomotor symptoms and irregular periods
   - menopause in women who have not had a period for at least 12 months
   - menopause based on symptoms in women without a uterus.

2. Take into account that it can be difficult to diagnose menopause in women taking sex steroids.

3. Do not use the following laboratory and imaging tests to diagnose perimenopause or menopause in women aged over 45 years:
   - anti-Müllerian hormone
   - inhibin A
   - inhibin B
   - oestradiol
   - antral follicle count
   - ovarian volume.

4. Do not use a serum follicle stimulating hormone (FSH) test to diagnose menopause in women using combined oestrogen and progestogen contraception or high-dose progestogen.

5. Consider using a FSH test to diagnose menopause only:
   - in women aged over 45 years with atypical symptoms
   - in women aged 40 to 45 years with menopausal symptoms, including a change in their menstrual cycle
   - in women aged under 40 years in whom menopause is suspected (see also section 11.1)

6. Give information to menopausal women and their family members or carers (as appropriate) that includes:
   - an explanation of the stages of menopause
   - common symptoms (see recommendation 8) and diagnosis
   - lifestyle changes and interventions that could help general health and wellbeing
   - the benefits and risks of treatments for menopausal symptoms.

7. Give information on menopause in different ways to help encourage women to discuss their symptoms and needs.

8. Explain to women that as well as a change in their menstrual cycle they may experience a variety of symptoms associated with menopause, including:
   - vasomotor symptoms (for example, hot flushes and sweats)
   - musculoskeletal symptoms (for example, joint and muscle pain)
   - effects on mood (for example, low mood)
   - urogenital symptoms (for example, vaginal dryness)
   - sexual difficulties (for example, low sexual desire).

9. Offer women who are likely to go through menopause as a result of medical or surgical treatment (including women with cancer, at high risk of hormone-sensitive cancer or having gynaecological surgery) support and:
• information about menopause and fertility before they have their treatment
• referral to a healthcare professional with expertise in menopause.

10. Adapt a woman’s treatment based on her changing symptoms as she goes through the stages of menopause.

11. Offer hormone replacement therapy (HRT) for vasomotor symptoms after discussing the short-term (up to 5 years) and longer-term benefits and risks. Offer a choice of oral or transdermal preparations as follows:
• oestrogen and progestogen to women with a uterus
• oestrogen alone to women without a uterus.

12. Do not routinely offer selective serotonin reuptake inhibitors (SSRIs) or serotonin and norepinephrine reuptake inhibitors (SNRIs) as first-line treatment for vasomotor symptoms alone.

13. Explain to women that although there is some evidence that isoflavones or black cohosh may relieve vasomotor symptoms, their safety is unknown and different preparations may vary.


15. Consider cognitive behavioural therapy (CBT) to alleviate low mood and anxiety in menopausal women.

16. Ensure that menopausal women and healthcare professionals involved in their care understand that there is no clear evidence for SSRIs or SNRIs to ease low mood in menopausal women who have not been diagnosed with depression (see the NICE guideline on depression in adults).

17. Consider testosterone\(^1\) supplementation for menopausal women with low sexual desire if HRT alone is not effective.

18. Offer low-dose vaginal oestrogen to women with urogenital atrophy (including those on systemic HRT) and continue treatment for as long as needed to relieve symptoms.

19. If systemic HRT is contraindicated, consider low-dose vaginal oestrogen after seeking advice from a healthcare professional with expertise in menopause.

20. If low-dose vaginal oestrogen does not relieve symptoms of urogenital atrophy, consider increasing the dose after seeking advice from a healthcare professional with expertise in menopause.

21. Explain to women with urogenital atrophy that:
• symptoms often come back when treatment is stopped
• adverse effects from low-dose vaginal oestrogen are very rare
• they should report unscheduled vaginal bleeding to their GP.

22. Advise women with vaginal dryness that moisturisers and lubricants can be used alone or in addition to vaginal oestrogen.

23. Do not offer routine monitoring of endometrial thickness during treatment for urogenital atrophy.

\(^1\) At the time of consultation (June 2015), testosterone did not have a UK marketing authorisation for this indication in women. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council’s prescribing guidance: prescribing unlicensed medicines for further information.
24. Explain to women that the efficacy and safety of unregulated compounded bioidentical hormones are unknown.

25. Explain to women who wish to try complementary therapies that the quality, purity and constituents of products may be unknown.

26. Explain to women with breast cancer that St John's wort may be a treatment option for menopausal symptoms but can interact with other medicines (for example, tamoxifen).

27. Discuss with women the importance of keeping up to date with nationally recommended health screening.

28. Review each treatment for short-term menopausal symptoms:
   - at 3 months to assess efficacy and tolerability
   - annually thereafter unless there are clinical indications for an earlier review (such as treatment ineffectiveness, side effects or adverse events).

29. Refer women to a healthcare professional with expertise in menopause if treatments do not improve their menopausal symptoms or they have ongoing troublesome side effects.

30. For women with menopausal symptoms and contraindications to HRT:
   - provide information on non-hormonal and non-pharmaceutical treatments (for example, CBT, hypnosis, acupuncture and relaxation techniques) for the relief of menopausal symptoms
   - consider referral to a healthcare professional with expertise in menopause.

31. Consider referring women to a healthcare professional with expertise in menopause if there is uncertainty about the most suitable treatment options for their menopausal symptoms.

32. Explain to women with a uterus that unscheduled vaginal bleeding is a common side effect of HRT within the first 3 months of treatment but should be reported at review appointments.

33. Offer women who are stopping HRT a choice of gradually reducing or immediately stopping treatment.

34. Explain to women that:
   - gradually reducing or immediately stopping HRT makes no difference to their symptoms in the longer term
   - gradually reducing HRT may limit recurrence of symptoms in the short term.

35. For advice on the treatment of menopausal symptoms in women with breast cancer or at high risk of breast cancer, see section 1.13 of the NICE guideline on early and locally advanced breast cancer and section 1.7 of the NICE guideline on familial breast cancer.

36. Offer menopausal women with or at high risk of breast cancer:
   - information on all available treatment options
   - referral to a healthcare professional with expertise in menopause.

37. Explain to women that:
   - the risk of venous thromboembolism (VTE) associated with HRT is greater for oral than transdermal preparations
   - the risk associated with transdermal HRT given at standard therapeutic doses is no greater than baseline risk.
38. Consider transdermal rather than oral HRT for menopausal women who are at increased risk of VTE, including those with a BMI over 30.

39. Refer menopausal women at high risk of VTE (for example, those with a strong family history of VTE or a hereditary thrombophilia) to a haematologist for assessment before considering HRT.

40. Ensure that menopausal women and healthcare professionals involved in their care understand that HRT:
   - does not increase cardiovascular disease risk when started in women aged under 60 years
   - does not affect the risk of dying from cardiovascular disease.

41. Be aware that cardiovascular risk factors (for example hypertension) do not automatically preclude a woman from taking HRT but should be taken into account.

42. Using tables 1 and 2, explain to women that:
   - the baseline risk of coronary heart disease and stroke for women around menopausal age varies from one woman to another according to the presence of cardiovascular risk factors
   - HRT with oestrogen alone is associated with no, or reduced, risk of coronary heart disease
   - HRT with oestrogen and progestogen is associated with little or no increase in the risk of coronary heart disease.

43. Explain to women that taking oral (but not transdermal) oestrogen is associated with a small increase in the risk of stroke. Also explain that the baseline risk of stroke in women aged under 60 years is very low (see table 2).

**Table 1: Absolute rates of CHD for different types of HRT compared with no HRT (or placebo), different duration of HRT use and time since stopping HRT for menopausal women**

<table>
<thead>
<tr>
<th></th>
<th>Past users</th>
<th>Current users</th>
<th>Treatment duration &lt;5 years</th>
<th>Treatment duration 5–10 years</th>
<th>&gt;5 years since stopping treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women on oestrogen alone</td>
<td>RCT estimate(^1)</td>
<td>–</td>
<td>7 fewer (from 11 fewer to 0)</td>
<td>–</td>
<td>6 fewer (from 9 fewer to 2 fewer)</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
<td>6 fewer (from 9 fewer to 3 fewer)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Women on oestrogen plus progestogen</td>
<td>RCT estimate(^1)</td>
<td>–</td>
<td>4 more (from 4 fewer to 17 more)</td>
<td>–</td>
<td>4 more (from 1 fewer to 11 more)</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Women on any HRT</td>
<td>RCT estimate</td>
<td>–</td>
<td>6 fewer (from 11 fewer to 5 more)</td>
<td>–</td>
<td>5 fewer (from 9 fewer to 3 more)</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>3 fewer (from 4 fewer to 1 fewer)</td>
<td>1 fewer (from 2 fewer to 0 fewer)</td>
<td>5 fewer (from 7 fewer to 3 fewer)</td>
<td>6 fewer (from 8 fewer to 4 fewer)</td>
<td>–</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial.
\(^1\) for women aged 50–59 years.
Table 2: Absolute rates of stroke for different types of HRT compared with no HRT (or placebo), different duration of HRT use and time since stopping HRT for menopausal women

<table>
<thead>
<tr>
<th>Difference in stroke incidence per 1000 menopausal women over 7.5 years (baseline risk in the UK population over 7.5 years: 11.3 women per 1000 [Weiner et al. 2008])</th>
<th>Past users</th>
<th>Current users</th>
<th>Treatment duration &lt;5 years</th>
<th>Treatment duration 5–10 years</th>
<th>&gt;5 years since stopping treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women on oestrogen alone</td>
<td>RCT estimate&lt;sup&gt;1&lt;/sup&gt;</td>
<td>–</td>
<td>1 more (from 5 fewer to 14 more)</td>
<td>–</td>
<td>1 more (from 4 fewer to 9 more)</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
<td>3 more (from 1 fewer to 8 more)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Women on oestrogen plus progestogen</td>
<td>RCT estimate&lt;sup&gt;1&lt;/sup&gt;</td>
<td>–</td>
<td>5 more (from 3 fewer to 20 more)</td>
<td>–</td>
<td>4 more (from 1 fewer to 13 more)</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
<td>4 more (from 1 more to 7 more)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Women on any HRT</td>
<td>RCT estimate</td>
<td>–</td>
<td>3 fewer (from 7 fewer to 8 more)</td>
<td>–</td>
<td>1 fewer (from 6 fewer to 7 more)</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>0 fewer (from 2 fewer to 2 more)</td>
<td>3 more (from 2 more to 5 more)</td>
<td>–</td>
<td>1 more (from 2 fewer to 4 more)</td>
<td>–</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial. <sup>1</sup>for women aged 50–59 years.

44. Explain to women that taking HRT (either orally or transdermally) is not associated with an increased risk of developing type 2 diabetes.

45. Ensure that women with type 2 diabetes and all healthcare professionals involved in their care are aware that HRT is not associated with an adverse effect on blood glucose control.

46. Consider HRT for menopausal symptoms in women with type 2 diabetes after taking comorbidities into account and seeking specialist advice if needed.

47. Ensure that menopausal women and healthcare professionals involved in their care understand that HRT does not affect the risk of dying from breast cancer.

48. Using table 3, explain to women around the age of natural menopause that:

- the baseline risk of breast cancer for women around menopausal age in the UK varies from one woman to another
- HRT with oestrogen alone is associated with little or no increase in the risk of breast cancer
- HRT with oestrogen and progestogen can be associated with an increase in the risk of breast cancer
- any increase in risk of breast cancer is related to treatment duration and reduces after stopping HRT.

Table 3: Absolute rates of breast cancer for different types of HRT compared with no HRT (or placebo), different duration of HRT use and time since stopping HRT for menopausal women

<table>
<thead>
<tr>
<th>Difference in breast cancer incidence per 1000 menopausal women (baseline risk in the UK population over 7.5 years: 9.45 women per 1000 [ONS data, 2010])</th>
<th>Past users</th>
<th>Current users</th>
<th>Treatment duration &lt;5 years</th>
<th>Treatment duration 5–10 years</th>
<th>&gt;5 years since stopping treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women on oestrogen alone</td>
<td>RCT estimate&lt;sup&gt;2&lt;/sup&gt;</td>
<td>–</td>
<td>1 more (from 5 fewer to 14 more)</td>
<td>1 fewer (from 4 fewer to 9 more)</td>
<td>1 fewer (from 3 fewer to 7 more)</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
<td>3 more (from 1 fewer to 8 more)</td>
<td>–</td>
<td>1 fewer (from 2 fewer to 4 more)</td>
<td>–</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial. <sup>2</sup>for women aged 50–59 years.
Menopause
Guideline summary

National Collaborating Centre for Women’s and Children’s Health

Difference in breast cancer incidence per 1000 menopausal women (baseline risk in the UK population over 7.5 years: 9.45 women per 1000 [ONS data, 2010])

<table>
<thead>
<tr>
<th></th>
<th>Past users</th>
<th>Current users</th>
<th>Treatment duration &lt;5 years</th>
<th>Treatment duration 5–10 years</th>
<th>&gt;5 years since stopping treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women on oestrogen alone</td>
<td>RCT estimate</td>
<td>–</td>
<td>3 fewer (from 6 fewer to 1 more)</td>
<td>–</td>
<td>2 fewer (from 5 fewer to 1 more)</td>
</tr>
<tr>
<td></td>
<td>Observational estimate</td>
<td>0 fewer (from 2 fewer to 3 more)</td>
<td>2 more (from 0 to 5 more)</td>
<td>4 more (from 0 to 5 more)</td>
<td>2 more (from 1 fewer to 6 more)</td>
</tr>
<tr>
<td>Women on oestrogen plus progestogen</td>
<td>RCT estimate</td>
<td>–</td>
<td>2 more (from 2 fewer to 8 more)</td>
<td>–</td>
<td>3 more (from 0 to 7 more)</td>
</tr>
<tr>
<td></td>
<td>Observational estimate</td>
<td>1 fewer (from 5 fewer to 5 more)</td>
<td>7 more (from 6 more to 8 more)</td>
<td>5 more (from 2 more to 8 more)</td>
<td>9 more (from 4 more to 16 more)</td>
</tr>
<tr>
<td>Women on any HRT</td>
<td>RCT estimate</td>
<td>–</td>
<td>4 fewer (from 7 fewer to 3 more)</td>
<td>–</td>
<td>1 fewer (from 5 fewer to 6 more)</td>
</tr>
<tr>
<td></td>
<td>Observational estimate</td>
<td>0 fewer (from 0 fewer to 1 more)</td>
<td>7 more (from 5 more to 10 more)</td>
<td>5 more (from 1 more to 9 more)</td>
<td>10 more (from 3 more to 19 more)</td>
</tr>
</tbody>
</table>

HRT, hormone replacement therapy; RCT, randomised controlled trial
1 For women aged 50–59 years

52. Explain to menopausal women that the likelihood of HRT affecting their risk of dementia is unknown.
53. Explain to women that:
   - there is limited evidence suggesting that HRT may improve muscle mass and strength
   - muscle mass and strength is maintained through, and is important for, activities of daily
     living.

54. Take into account the woman’s clinical history (for example, previous medical or surgical
    treatment) and family history when diagnosing premature ovarian insufficiency.

55. Diagnose premature ovarian insufficiency in women aged under 40 years based on:
   - menopausal symptoms, including no or infrequent periods (taking into account whether
     the woman has a uterus) and
   - elevated FSH levels on 2 blood samples taken 4–6 weeks apart.

56. Do not diagnose premature ovarian insufficiency on the basis of a single blood test.

57. Do not routinely use anti-Müllerian hormone testing to diagnose premature ovarian
    insufficiency.

58. If there is doubt about the diagnosis of premature ovarian insufficiency, consider anti-
    Müllerian hormone testing after seeking specialist advice (see the NICE guideline on fertility).

59. Offer sex steroid replacement with a choice of HRT or a combined oral contraceptive to
    women with premature ovarian insufficiency, unless contraindicated (for example, in women
    with hormone-sensitive cancer).

60. Explain to women with premature ovarian insufficiency:
   - the importance of starting hormonal treatment either with HRT or a combined oral
     contraceptive and continuing treatment until at least the age of natural menopause (unless
     contraindicated).
   - that HRT may have a beneficial effect on blood pressure when compared with a combined
     oral contraceptive
   - that both HRT and combined oral contraceptives offer bone protection
   - that they should not use HRT as a contraceptive.

61. Give women with premature ovarian insufficiency and contraindications to hormonal
    treatments advice on bone and cardiovascular health, and symptom management (see also
    section 7).

1.4 Research recommendations

1. What is the efficacy of different treatments for menopausal symptoms in women who
   have had treatment for, or are at risk of, breast cancer?

2. What is the impact of systemic HRT usage in women with a previous diagnosis of breast
   cancer for the risk of breast cancer recurrence, mortality or tumour aggression?

3. How does the preparation of HRT affect the risk of venous thromboembolism (VTE)?

4. What is the difference in the risk of breast cancer in menopausal women on HRT with
   either progesterone, progestogen or selective oestrogen receptor modulators?

5. What is the impact of oestradiol in combination with the levonorgestrel-secreting intra-
   uterine system (LNG-IUS) on the risk of breast cancer and venous thrombo-embolism
   (VTE)?

6. What are the effects of early HRT use on the risk of dementia?

7. What are the main clinical manifestations of premature ovarian insufficiency and the
   short- and long-term impact of the most common therapeutic interventions?
1.5 Other versions of the guideline

Details about the other versions of the guideline (such as the NICE pathway and the Information for the Public) will be inserted here in the final published guideline.

1.6 Schedule for updating the guideline

For the most up-to-date information about the guideline reviews, please see the latest version of the NICE guidelines manual available from the NICE website www.nice.org.uk.
2 Development of the guideline

2.1 What is a NICE clinical guideline?

NICE clinical guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the NHS – from prevention and self-care through primary and secondary care to more specialised services. We base our clinical guidelines on the best available research evidence, with the aim of improving the quality of health care. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific review questions.

NICE clinical guidelines can:
- provide recommendations for the treatment and care of people by health professionals
- be used to develop standards to assess the clinical practice of individual health professionals
- be used in the education and training of health professionals
- help patients to make informed decisions
- improve communication between patient and health professional.

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:
- Guideline topic is referred to NICE from the Department of Health.
- Stakeholders register an interest in the guideline and are consulted throughout the development process.
- The scope is prepared by the National Collaborating Centre for Women and Children’s Health (NCC-WCH).
- The NCC-WCH establishes a guideline development group (GDG).
- A draft guideline is produced after the group assesses the available evidence and makes recommendations.
- There is a consultation on the draft guideline.
- The final guideline is produced.

The NCC-WCH and NICE produce a number of versions of this guideline:
- the ‘full guideline’ contains all the recommendations, together with details of the methods used and the underpinning evidence
- the ‘NICE guideline’ lists the recommendations
- ‘information for the public’ is written using suitable language for people without specialist medical knowledge
- NICE Pathways brings together all connected NICE guidance.

2.2 Remit

NICE received the remit for this guideline from the Department of Health. They commissioned the NCC-WCH to produce the guideline.

The remit for this guideline is to develop a clinical guideline on the diagnosis and management of menopause.
2.3 Who developed this guideline?

A multidisciplinary Guideline Development Group (GDG) comprising health professionals and researchers as well as lay members developed this guideline (see the list of Guideline Development Group members and acknowledgements).

The National Institute for Health and Care Excellence (NICE) funds the National Collaborating Centre for Women and Children’s Health (NCC-WCH) and thus supported the development of this guideline. The GDG was convened by the NCC-WCH and chaired by Professor Mary Ann Lumsden in accordance with guidance from NICE.

The group met every 4-6 weeks during the development of the guideline. At the start of the guideline development process all GDG members declared interests including consultancies, fee-paid work, share-holdings, fellowships and support from the healthcare industry. At all subsequent GDG meetings, members declared arising conflicts of interest.

Members were either required to withdraw completely or for part of the discussion if their declared interest made it appropriate. The details of declared interests and the actions taken are shown in Appendix C:

Staff from the NCC-WCH provided methodological support and guidance for the development process. The team working on the guideline included a project manager, systematic reviewers, health economists and information scientists. They undertook systematic searches of the literature, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate and drafted the guideline in collaboration with the GDG.

2.4 What this guideline covers

2.4.1 Groups that will be covered

- menopausal women (covering the perimenopause and postmenopause)
- women with premature ovarian insufficiency (irrespective of cause)

2.4.2 Key clinical issues that will be covered

- diagnosis and classification of the stages of menopause
- optimal clinical management of menopause-related symptoms, including:
  - treatments for symptomatic relief (specifically vasomotor, musculoskeletal and psychological symptoms, and altered sexual function), including:
    - hormonal pharmaceutical treatments:
      - oestrogen combined with progestogen (oral)
      - oestrogen combined with progestogen (transdermal)
      - oestrogen (oral)
      - oestrogen (transdermal)
      - oestrogen (depot)
      - progestogen alone
      - testosterone
      - tibolone
      - bio-identical hormones licensed for use in the UK
      - tissue-selective oestrogen complexes
      - selective oestrogen-receptor modulators
    - non-hormonal pharmaceutical treatments:
- selective serotonin reuptake inhibitors
- serotonin–noradrenaline reuptake inhibitors
- gabapentin
- clonidine
- non-pharmaceutical treatments:
  - phytoestrogens
  - herbal preparations (including black cohosh and red clover)
  - acupuncture
  - lifestyle advice
- psychological therapies
  - cognitive behavioural therapy
- risks and benefits of treatments
- timing of treatment
- monitoring of treatment
- duration of treatment
- treatment withdrawal strategies

Note that guideline recommendations will normally fall within licensed indications. Exceptionally, and only if clearly supported by evidence, use outside a licensed indication may be recommended. The guideline will assume that prescribers will use a drug’s summary of product characteristics to inform decisions made with individual patients.

- contribution of HRT in preventing long-term sequelae of the menopause (especially osteoporosis and CVD)
- diagnosis and management of premature ovarian insufficiency

For further details please refer to the scope in Appendix A and review questions in Appendix D.

2.5 What this guideline does not cover

2.5.1 Groups that will not be covered
- women who are pregnant
- women who are breastfeeding
- men
- transgender women

2.5.2 Clinical issues that will not be covered
- contribution of all other agents (excluding HRT) in preventing long-term sequelae of the menopause
- systemic oestrogen-based hormonal treatment in women who have an increased risk of, or are undergoing treatment for, breast cancer
- treatment and/or prevention of chronic diseases that are common in post-menopausal women e.g. osteoporosis and CVD
- premenopausal prevention of symptoms usually associated with the menopause (specifically vasomotor, musculoskeletal, urogenital and psychological symptoms and altered sexual function)
- investigation of the cause of premature ovarian insufficiency in women presenting with primary amenorrhea
Menopause

Development of the guideline

- induction of puberty in children and young people
- cost-effectiveness analysis of methods of contraception during the menopause

2.6 Relationships between the guideline and other NICE guidance

2.6.1 Related NICE guidance

- Osteoporosis (2012). NICE clinical guideline 146.
- Epilepsy (2012). NICE clinical guideline 137.
- Denosumab for the prevention of osteoporotic fractures in postmenopausal women (2010). NICE technology appraisal 204.
- Heavy menstrual bleeding (2007). NICE clinical guideline 44.
- Patient experience in adult NHS services (2012). NICE guideline CG138
- Medicines adherence (2009). NICE guideline CG76
- Fertility (2013). NICE clinical guideline 156.
3 Guideline development methodology

This chapter sets out in detail the methods used to review the evidence and to generate the recommendations that are presented in subsequent chapters. This guidance was developed in accordance with the methods outlined in the NICE guidelines manual 2012 for the stages up to guideline development and moved to the updated NICE guidelines manual 2014 since consultation stage.

3.1 Developing the review questions and protocols

Review questions were developed in a PICO framework (patient, intervention, comparison and outcome) for intervention reviews, in a framework of population, index tests, reference standard and target condition for reviews of diagnostic test accuracy and using population, area of interest and outcomes for qualitative reviews.

This use of a framework guided the literature searching process, critical appraisal and synthesis of evidence and facilitated the development of recommendations by the Guideline Development Group (GDG). The review questions were drafted by the NCC-WCH technical team and refined and validated by the GDG. The questions were based on the key clinical areas identified in the scope (Appendix A).

A total of 17 review questions were identified (Table 5).

Full literature searches, critical appraisals and evidence reviews were completed for all the specified review questions.

Table 5: Description of review questions

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Type of review</th>
<th>Review questions</th>
<th>Outcomes</th>
</tr>
</thead>
</table>
| 4       | Diagnostic review | What is the diagnostic accuracy of the following indicators (clinical and biological manifestations) in the diagnosis of perimenopause and postmenopause:  
• age  
• menopausal symptoms (especially vasomotor symptoms)  
• endocrine changes (specifically follicle-stimulating hormone (FSH), anti-Müllerian hormone, oestrogen or inhibin B)  
• total antral follicle count (AFC) | • sensitivity / specificity  
• likelihood ratio (positive and negative)  
• area under the curve (AUC) |
| 5       | Comparative review | What is the usefulness of formal classification systems compared with non-structured classification systems in the diagnosis of menopause and in guiding further treatment? | • correct diagnosis of menopause  
• guidance for further investigation or treatment  
• HRQoL |
| 6       | Qualitative review | What are the information needs for women in menopause? | 1st part of question:  
• areas of information need  
2nd part of question:  
• woman’s knowledge about menopause  
• number of visits to the health care professionals |
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Type of review</th>
<th>Review questions</th>
<th>Outcomes regarding menopause issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Interventional review</td>
<td>What is the most clinical and cost effective treatment for the relief of individual menopause-related symptoms for women in menopause?</td>
<td>frequency of hot flushes (including night sweats) frequency of sexual activity psychological symptoms anxiety low mood (not clinical depression) musculoskeletal symptoms safety outcomes discontinuation vaginal bleeding</td>
</tr>
<tr>
<td>7.11</td>
<td>Interventional review</td>
<td>What is the clinical effectiveness of local oestrogens and ospemefine compared with placebo for menopause-related vaginal/urogenital atrophy?</td>
<td>Efficacy outcomes: measurement of vaginal pH, maturation index patient assessment of symptoms improvement itching and discomfort Safety outcomes: assessment of endometrial stimulation breast pain (a surrogate marker for systemic absorption, and blood oestradiol levels) frequency of adverse events relating to treatment acceptability withdrawal from the study because of adverse events relating to treatment participant adherence to treatment health-related quality of life Long term outcomes endometrial hyperplasia or cancer confirmed by biopsy symptom relief health-related quality of life</td>
</tr>
<tr>
<td>8</td>
<td>Interventional review</td>
<td>At what intervals should clinical review be undertaken to assess the effectiveness and safety of treatments to relieve menopausal symptoms and to determine when women need to be</td>
<td>reoccurrence of menopausal symptoms HRQoL resumption of HRT</td>
</tr>
<tr>
<td>Chapter</td>
<td>Type of review</td>
<td>Review questions</td>
<td>Outcomes</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>------------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>referred to specialist care?</td>
<td>treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• uptake of alternative treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• acceptability of treatment to women (qualitative assessment if scale not available)</td>
</tr>
<tr>
<td>33</td>
<td>Interventional review</td>
<td>In perimenopausal and postmenopausal women using HRT for vasomotor symptom relief, what is the clinical effectiveness of an abrupt HRT discontinuation strategy compared with a tapered HRT discontinuation strategy?</td>
<td>• reoccurrence of menopausal symptoms</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• HRQoL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• resumption of HRT treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• uptake of alternative treatment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• acceptability of treatment to women (qualitative assessment if scale not available)</td>
</tr>
<tr>
<td>10.1</td>
<td>Interventional review</td>
<td>What are the effects of HRT administered for menopausal symptoms on the risk of developing VTE?</td>
<td>• VTE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• mortality (overall or included condition specific mortality)</td>
</tr>
<tr>
<td>10.2</td>
<td>Interventional review</td>
<td>What are the effects of the risk of HRT administered for menopausal symptoms on the risk of development of CVD (including stroke)</td>
<td>• change in blood pressure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• stroke</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• myocardial infarction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• cardiac event composite scores</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• mortality – cardio related</td>
</tr>
<tr>
<td>10.3</td>
<td>Interventional review</td>
<td>What are the effects of HRT administered for menopausal symptoms on the risk of developing T2DM?</td>
<td>• T2DM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• mortality (overall or included condition specific mortality)</td>
</tr>
<tr>
<td>44</td>
<td>Interventional review</td>
<td>What impact does use of HRT for menopausal symptoms have on control of diabetes/glycaemic levels in those with T2DM?</td>
<td>• HbA1c</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• hyperglycaemic episodes (self-monitoring, finger prick tests)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• HRQoL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• mortality (overall or included condition specific mortality)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• adverse effects (complications resulting from diabetes)</td>
</tr>
<tr>
<td>46</td>
<td>Interventional review</td>
<td>What are the effects of HRT administered for menopausal symptoms on risk of developing breast cancer?</td>
<td>• breast cancer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• mortality from breast cancer</td>
</tr>
<tr>
<td>10.6</td>
<td>Interventional review</td>
<td>What are the effects of HRT administered for menopausal symptoms on the risk of development of osteoporosis?</td>
<td>• vertebral fracture</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• neck of femur (hip) fracture</td>
</tr>
</tbody>
</table>
### 3.2 Searching for evidence

#### 3.2.1 Clinical literature search

Systematic literature searches were undertaken to identify all published clinical evidence relevant to the review questions.

Databases were searched using relevant medical subject headings, free-text terms and study type filters where appropriate. Studies published in languages other than English were not reviewed. Where possible, searches were restricted to retrieve articles published in English. All searches were conducted in MEDLINE, Embase, and The Cochrane Library. All searches were updated on 22nd January 2015. Due to the complexity of the NMA and the time implications of updating the data analysis, searches were updated at an earlier date, on 13th January 2015. Any studies added to the databases after this date (even those published prior to this date) were not included unless specifically stated in the text.

Search strategies were quality assured by cross checking reference lists of highly relevant papers, analysing search strategies in other systematic reviews and asking the GDG.
members to highlight any additional studies. The questions, the study types applied, the
databases searched and the years covered can be found in Appendix E.

The titles and abstracts of records retrieved by the searches were sifted for relevance, with
potentially significant publications obtained in full text. These were assessed against the
inclusion criteria.

During the scoping stage, a search was conducted for guidelines and reports on websites of
organisations relevant to the topic. Searching for grey literature or unpublished literature was
not undertaken. Searches for electronic, ahead of print publications are not routinely
undertaken unless indicated by the GDG. All references suggested by stakeholders at the
scoping consultation were initially considered.

3.3 Reviewing and synthesising the evidence

The evidence was reviewed following the steps:

- potentially relevant studies were identified for each review question from the relevant
  search results by reviewing titles and abstracts. Full papers were then obtained
- full papers were reviewed against pre-specified inclusion and exclusion criteria in the
  review protocols (in Appendix D) and were presented in summary tables in each chapter
  and evidence tables (in Appendix H)
- relevant studies were critically appraised using the appropriate checklist as specified in
  the guidelines manual 2012
- summaries of evidence were generated by outcome and were presented in GDG
  meetings:
  o randomised studies: data were meta-analysed where appropriate and reported in
    GRADE profiles (for interventional reviews).
  o observational studies: data were presented as a range of values or meta-analysed
    (where appropriate) in GRADE profiles and usually this was organised by outcomes.
  o diagnostic accuracy studies were presented as measures of diagnostic test accuracy
    (sensitivity, specificity, positive and negative likelihood ratio, area under the curve) in a
    modified version of a GRADE profile. A meta-analysis was not conducted when
    included studies were too heterogeneous
  o qualitative studies: the themes of the studies were organised in summary evidence
    tables, along with quality assessment otherwise presented in a narrative form
- 80% of all data extracted was quality assured by a second reviewer. 50% of the GRADE
  quality assessment was quality assured by a second reviewer to minimise any potential
  risk of reviewer bias or error.

3.3.1 Methods of combining clinical studies

3.3.1.1 Data synthesis for intervention reviews

Where possible, meta-analyses were conducted to combine the results of studies for each
review question using Cochrane Review Manager (RevMan5) software or STATA. Fixed-
effects (Mantel-Haenszel) techniques were used to calculate risk ratios (relative risk) for the
binary outcomes.

For the continuous outcomes, measures of central tendency (mean) and variation (standard
deviation) were required for meta-analysis. A generic inverse variance option in RevMan5
was used if any studies reported solely the summary statistics and 95% confidence interval
(95% CI) or standard error; this included any hazard ratios reported. However, in cases
where standard deviations were not reported per intervention group, the standard error (SE)
for the mean difference was calculated from other reported statistics (p values or 95% CIs) if
available; meta-analysis was then undertaken for the mean difference and SE using the
generic inverse variance method in RevMan5. When the only evidence was based on studies
that summarised results by presenting medians (and interquartile ranges), or only p values
were given, this information was assessed in terms of the study’s sample size and was
included in the GRADE tables as a narrative summary. Consequently, aspects of quality
assessment such as imprecision of effect could not be assessed for this evidence and this
has been recorded in the footnotes of the GRADE tables.

In instances where multiple scales were reported for a single outcome, mean differences
were standardised (divided by their SD) before pooling, giving meta-analysed results that
were reported as standardised mean differences (SMD), with a standard deviation of 1.

Where reported, time-to-event data were presented as a hazard ratio or results from a Cox
hazard proportion model were given as a result from a multivariate analysis.

Stratified analyses were predefined for some review questions at the protocol stage when the
GDG identified that these strata to be different in terms of clinical characteristics and the
interventions were expected to have a different effect, for example on the management of
short term symptoms. We stratified our analysis for women with uterus, without uterus and
women with a history of breast cancer.

Statistical heterogeneity was assessed by visually examining the forest plots, and by
considering the chi-squared test for significance at p < 0.1 or an I-squared inconsistency
statistic (with an I-squared value of 50-74.99% indicating serious inconsistency and I-
squared value of over 75% indicating very serious inconsistency). If the heterogeneity still
remained, a random-effects (DerSimonian and Laird) model was employed to provide a more
conservative estimate of the effect. Where considerable heterogeneity was present, we set
to out to perform predefined subgroup analyses based on the following factors:

• different stages of menopause (peri or postmenopausal)
• different age groups (below 45 years old, over 50 or 60 years)

3.3.1.2 Data synthesis for diagnostic test accuracy review

For diagnostic test accuracy studies, the following outcomes were reported: sensitivity,
specificity, positive and negative likelihood ratio and area under the curve (AUC)

3.3.1.3 Data synthesis for qualitative review

For the qualitative review in the guideline, results were reported narratively either by
individual study or by summarising the range of values as reported across similar studies. A
summary evidence table was used when data allowed for this.

3.3.1.4 Data synthesis using network meta-analysis (NMA)

A NMA was formulated to synthesise direct and indirect evidence of treatments’ efficacy to
relieve short term menopausal symptoms whilst preserving randomisation for the outcomes
of frequency of vasomotor symptoms, discontinuation of treatment, and vaginal bleeding.
Hierarchical Bayesian network meta-analyses (NMAs) with class effects was performed
using the software WinBUGS version 1.4. Data from women in 3 distinct populations were
used as inputs to the models – women with a uterus, women without a uterus, and women
with breast cancer or a history of breast cancer. We examined statistical models for fixed and
random effects that allowed inclusion of multi arm trials and accounts for the correlation
between arms in the trials with any number of trial arms. These models was based on
original work from the University of Bristol

(https://www.bris.ac.uk/cobm/research/mpes/mtc.html).
As no dependency on time was identified, discontinuation of treatment and vaginal bleeding were treated as dichotomous outcomes and were modelled on the log-odds ratio scale. Frequency of vasomotor symptoms was distributed in the form of an overdispersed Poisson distribution and was therefore modelled on the log-mean ratio scale. On this scale, final and change from baseline frequencies of vasomotor symptoms could not be pooled so a correlation coefficient was used to estimate final frequencies from change from baseline.

For all the networks set up in the NMA, models for fixed and random effects were developed and then these were compared based on residual deviance and deviance information criteria (DIC). The model with the smallest DIC is estimated to be the model that would best predict a replicate dataset which has the same structure as that currently observed. A small difference in DIC between the fixed and random effects models (3-5 points) implies that the better fit obtained by adding random effects does not justify the additional complexity. However, if the difference in DIC between a fixed and random effect model was less than 5 points, and the models make very similar inferences, then we would report the results from a fixed effects model results as it does not make as many assumptions as the random effect model, contains fewer parameters and it is easier for clinical interpretation than the random effects model.

Where closed loops of treatment comparisons existed in the networks, inconsistency was assessed by comparing any available direct and indirect treatment comparison and testing the null hypothesis that the indirect evidence was no different than the direct evidence.

There were 3 main outputs from the NMA: 1) the estimation of summary estimates (MRs or ORs) (with their 95% credible intervals) were calculated for comparisons of the direct and indirect evidence, 2) the probability that each treatment was best based on the proportion of Markov chain iterations in which treatment had the highest probability of achieving the outcomes selected in the networks and 3) the ranking of treatments compared to baseline groups (presented as median rank and its 95% credible intervals).

The following sensitivity analyses were conducted:
- changes to the value of the correlation coefficient used to estimate final frequencies of vasomotor symptoms from change from baseline
- combining women with and without uterus into a single population to determine if this led to changes in heterogeneity
- removing low dose oral oestradiol plus progestogen to determine if this dose was reducing the overall efficacy of oral oestradiol plus progestogen in the model

### 3.3.2 Type of studies

RCTs, non-randomised trials, and observational studies (including diagnostic or comparative cohorts) were included in the evidence reviews as appropriate.

Literature reviews, posters, letters, editorials, comment articles, unpublished studies and studies not in English were excluded.

For most intervention reviews in this guideline, parallel RCTs were included because they are considered the most robust study design for unbiased estimation of intervention effects. Crossover RCTs were appropriate for some of the interventional questions.

If there was limited evidence from RCTs, well-conducted non-randomised comparative studies were included. For most review questions investigating long term outcomes of HRT, prospective comparative studies with adjusted analyses on important confounders were selected in addition to RCTs. Please refer to Appendix D for full details on the study design of studies selected for each review question.
For diagnostic reviews, cross-sectional and retrospective studies were included. Case-control or case series were not included for the presentation of evidence for any review question.

### 3.3.3 Appraising the quality of evidence by outcomes

The evidence for outcomes from the included RCTs and, where appropriate, observational studies was evaluated and presented using an adaptation of the ‘Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox’ developed by the international GRADE working group (http://www.gradeworkinggroup.org/). The software developed by the GRADE working group (GRADEpro) was used to assess the quality of each outcome, taking into account individual study quality factors and the meta-analysis results. The ‘Clinical/Economic evidence profile’ table includes details of the quality assessment and pooled outcome data, where appropriate, an absolute measure of intervention effect and the summary of quality of evidence for that outcome. In this table, the columns for intervention and control indicate summary measures and measures of dispersion (such as mean and standard deviation or median and range) for continuous outcomes and frequency of events (n/N: the sum across studies of the number of patients with events divided by the number of completers) for binary outcomes. Reporting or publication bias was only taken into consideration in the quality assessment and included in the ‘Clinical evidence profile’ table if it was apparent.

The selection of outcomes for each review question was decided when each review protocol was discussed with the GDG. However, given the nature of most of the review questions included in this guideline development (short or long term outcomes driven), the categorisation of outcomes as critical and important did not follow the standard GRADE approach. The outcomes selected for a review question were critical for decision making in a specific context.

The evidence for each outcome in interventional reviews was examined separately for the quality elements listed and defined in Table 6. Each element was graded using the quality levels listed in Table 7.

The main criteria considered in the rating of these elements are discussed below. Footnotes were used to describe reasons for grading a quality element as having serious or very serious limitations. The ratings for each component were summed to obtain an overall assessment for each outcome (Table 8).

The GRADE toolbox is currently designed only for RCTs and observational studies but we adapted the quality assessment elements and outcome presentation for diagnostic accuracy and qualitative studies subject to data availability.

For example, for diagnostic accuracy studies, the GRADE tables were modified to include the most appropriate measures of diagnostic accuracy (sensitivity, specificity, positive and negative likelihood ratio) whereas qualitative studies were presented in summary evidence tables around themes identified or direct participants’ quotations along a quality assessment per study level.

### Table 6: Description of quality elements in GRADE for intervention studies

<table>
<thead>
<tr>
<th>Description of quality elements in GRADE for intervention studies</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of bias (study limitations)</td>
<td>Limitations in the study design and implementation may bias the estimates of the treatment effect. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect.</td>
</tr>
</tbody>
</table>
Inconsistency refers to an unexplained heterogeneity of results.

Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question, or recommendation made, such that the effect estimate is changed.

Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. Imprecision results if the confidence interval includes the clinically important threshold.

Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.

<table>
<thead>
<tr>
<th>Inconsistency</th>
<th>Inconsistency refers to an unexplained heterogeneity of results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirectness</td>
<td>Indirectness refers to differences in study population, intervention, comparator and outcomes between the available evidence and the review question, or recommendation made, such that the effect estimate is changed.</td>
</tr>
<tr>
<td>Imprecision</td>
<td>Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. Imprecision results if the confidence interval includes the clinically important threshold.</td>
</tr>
<tr>
<td>Publication bias</td>
<td>Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies.</td>
</tr>
</tbody>
</table>

Table 7: Levels of quality elements in GRADE Level

<table>
<thead>
<tr>
<th>Levels of quality elements in GRADE Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>There are no serious issues with the evidence.</td>
</tr>
<tr>
<td>Serious</td>
<td>The issues are serious enough to downgrade the outcome evidence by one level.</td>
</tr>
<tr>
<td>Very serious</td>
<td>The issues are serious enough to downgrade the outcome evidence by 2 levels.</td>
</tr>
</tbody>
</table>

Table 8: Overall quality of outcome evidence in GRADE Level

<table>
<thead>
<tr>
<th>Table 4: Overall quality of outcome evidence in GRADE Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Further research is very unlikely to change our confidence in the estimate of effect.</td>
</tr>
<tr>
<td>Moderate</td>
<td>Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.</td>
</tr>
<tr>
<td>Low</td>
<td>Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.</td>
</tr>
<tr>
<td>Very low</td>
<td>Any estimate of effect is very uncertain.</td>
</tr>
</tbody>
</table>

3.3.3.1 Grading the quality of clinical evidence

After results were pooled, the overall quality of evidence for each outcome was considered. The following procedure was adopted when using GRADE approach:

1. A quality rating was assigned, based on the study design. RCTs start high, observational studies as moderate, and uncontrolled case series as low or very low.

2. The rating was then downgraded for the specified criteria: risk of bias (study limitations), inconsistency, indirectness, imprecision and publication bias. These criteria are detailed below. Evidence from observational studies (which had not previously been downgraded) was upgraded if there was: a large magnitude of effect, a dose-response gradient, and if all plausible confounding would reduce a demonstrated effect or suggest a spurious effect when results showed no effect. Each quality element considered to have 'serious' or 'very serious' risk of bias was rated down by 1 or 2 points respectively.
3. The downgraded/upgraded ratings were then summed and the overall quality rating was revised. For example, all RCTs started as High and the overall quality became Moderate, Low or Very low if 1, 2 or 3 points were deducted respectively.

4. The reasons or criteria used for downgrading were specified in the footnotes.

The details of the criteria used for each of the main quality elements are discussed further in the following sections (3.3.4.2 to 3.3.4.6).

3.3.2 Risk of bias

Bias can be defined as anything that causes a consistent deviation from the truth. Bias can be perceived as a systematic error, for example, if a study was carried out several times and there was a consistently wrong answer, the results would be inaccurate.

The risk of bias for a given study and outcome is associated with the risk of over- or underestimation of the true effect.

The risks of bias are listed in Table 9.

A study with a poor methodological design does not automatically imply high risk of bias; the bias is considered individually for each outcome and it is assessed whether this poor design will impact on the estimation of the intervention effect.

Table 9: Risk of bias in randomised controlled trials

<table>
<thead>
<tr>
<th>Risk of bias</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allocation concealment</td>
<td>Those enrolling patients are aware of the group to which the next enrolled patient will be allocated (this is a major problem in “pseudo” or “quasi” randomised trials with allocation by for example, day of week, birth date, chart number).</td>
</tr>
<tr>
<td>Lack of blinding</td>
<td>Patient, caregivers, those recording outcomes, those adjudicating outcomes, or data analysts are aware of the arm to which patients are allocated.</td>
</tr>
<tr>
<td>Incomplete accounting of patients and outcome events</td>
<td>Missing data not accounted for and failure of the trialists to adhere to the intention to treat principle when indicated.</td>
</tr>
<tr>
<td>Selective outcome reporting</td>
<td>Reporting of some outcomes and not others on the basis of the results.</td>
</tr>
<tr>
<td>Other risks of bias</td>
<td>For example:</td>
</tr>
<tr>
<td></td>
<td>• stopping early for benefit observed in randomised trials, in particular in the absence of adequate stopping rules</td>
</tr>
<tr>
<td></td>
<td>• use of unvalidated patient-reported outcomes</td>
</tr>
<tr>
<td></td>
<td>• recruitment bias in cluster randomised trials.</td>
</tr>
</tbody>
</table>

3.3.3 Diagnostic studies

For diagnostic accuracy studies, the Quality Assessment of Diagnostic Accuracy Studies version 2 (QUADAS-2) checklist was used. Risk of bias and applicability in primary diagnostic accuracy studies in QUADAS-2 consists of 4 domains (see Figure 2):

- patient selection
- index test
- reference standard
- flow and timing.
Figure 2: Summary of QUADAS-2 with a reference to quality domains

<table>
<thead>
<tr>
<th>DOMAIN</th>
<th>PATIENT SELECTION</th>
<th>INDEX TEST</th>
<th>REFERENCE STANDARD</th>
<th>FLOW AND TIMING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting).</td>
<td>Describe the index test and how it was conducted and interpreted.</td>
<td>Describe the reference standard and how it was conducted and interpreted.</td>
<td>Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard.</td>
</tr>
<tr>
<td>Signalling questions (yes/no/unclear)</td>
<td>Was a consecutive or random sample of patients enrolled?</td>
<td>Were the index test results interpreted without knowledge of the results of the reference standard?</td>
<td>Is the reference standard likely to correctly classify the target condition?</td>
<td>Was there an appropriate interval between index test(s) and reference standard?</td>
</tr>
<tr>
<td>Risk of bias: High/low/unclear</td>
<td>Could the selection of patients have introduced bias?</td>
<td>Could the conduct or interpretation of the index test have introduced bias?</td>
<td>Could the reference standard, its conduct, or its interpretation have introduced bias?</td>
<td>Could the patient flow have introduced bias?</td>
</tr>
<tr>
<td>Concerns regarding applicability: High/low/unclear</td>
<td>Are there concerns that the included patients do not match the review question?</td>
<td>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</td>
<td>Are there concerns that the target condition as defined by the reference standard does not match the review question?</td>
<td></td>
</tr>
</tbody>
</table>

3.3.3.4 Inconsistency

Inconsistency refers to an unexplained heterogeneity of results. When estimates of the treatment effect across studies differ widely (that is when there is heterogeneity or variability in results), this suggests true differences in underlying treatment effect.

Heterogeneity in meta-analyses was examined and sensitivity and subgroup analyses performed as pre-specified in the protocols (Appendix D).

When heterogeneity existed (chi-squared $p < 0.1$, I-squared inconsistency statistic of between 50-74.99% or I-squared > 50% or evidence from examining forest plots), but no plausible explanation was found (for example duration of intervention or different follow-up periods) the quality of evidence was downgraded by 1 or 2 levels, depending on the extent of uncertainty to the results contributed by the inconsistency in the results. In addition to the I-squared and chi-squared values, the decision for downgrading was also dependent on factors such as whether the intervention is associated with benefit in all other outcomes or whether the uncertainty about the magnitude of benefit (or harm) of the outcome showing heterogeneity would influence the overall judgment about net benefit or harm (across all outcomes).

When outcomes are derived from a single trial, inconsistency is not an issue for downgrading the quality of evidence. However, “no inconsistency” is nevertheless used to describe this quality assessment in the GRADE tables.

3.3.3.5 Indirectness

Directness refers to the extent to which the populations, intervention, comparisons and outcome measures are similar to those defined in the inclusion criteria for the reviews. Indirectness is important when these differences are expected to contribute to a difference in effect size, or may affect the balance of harms and benefits considered for an intervention.
3.3.3.6 Imprecision

Imprecision in guideline development concerns whether the uncertainty (confidence interval) around the effect estimate means that it is not clear whether there is a clinically important difference between interventions or not. Therefore, imprecision differs from the other aspects of evidence quality in that it is not really concerned with whether the point estimate is accurate or correct (has internal or external validity) instead it is concerned with the uncertainty about what the point estimate is. This uncertainty is reflected in the width of the confidence interval.

The 95% confidence interval (95% CI) is defined as the range of values that contain the population value with 95% probability. The larger the trial, the smaller the 95% CI and the more certain the effect estimate.

Imprecision in the evidence reviews was assessed by considering whether the width of the 95% CI of the effect estimate was relevant to decision-making, considering each outcome in isolation.

When the confidence interval of the effect estimate is wholly contained in one of the 3 zones (clinically important benefit, clinically important harm, no clinically important benefit or harm) we are not uncertain about the size and direction of effect (whether there is a clinically important benefit, or the effect is not clinically important, or there is a clinically important harm), so there is no imprecision.

When a wide confidence interval lies partly in each of 2 zones, it is uncertain in which zone the true value of effect estimate lies, and therefore there is uncertainty over which decision to make (based on this outcome alone). The confidence interval is consistent with 2 decisions and so this is considered to be imprecise in the GRADE analysis and the evidence is downgraded by 1 level (‘serious imprecision’).

If the confidence interval of the effect estimate crosses into 3 zones, this is considered to be very imprecise evidence because the confidence interval is consistent with 3 clinical decisions and there is a considerable lack of confidence in the results. The evidence is therefore downgraded by 2 levels in the GRADE analysis (‘very serious imprecision’).

Implicitly, assessing whether the confidence interval is in, or partially in, a clinically important zone, requires the GDG to estimate an MID or to say whether they would make different decisions for the 2 confidence limits.

Originally, the GDG was asked about MIDs in the literature or well established MIDs in the clinical community (for example international consensus documents) for the relevant outcomes of interest. Due to lack of well-established and widely accepted MIDs in the literature around menopause, the GDG agreed to use the GRADE default MIDs.

Therefore, the GDG considered it clinically acceptable to use the GRADE default MID to assess imprecision: a 25% relative risk reduction or relative risk increase was used, which corresponds to clinically important thresholds for a risk ratio of 0.75 and 1.25 respectively. This default MID was used for all the dichotomous outcomes in the interventions evidence reviews. For continuous outcomes, a MID was calculated by adding or subtracting 0.5 times standard deviations. For outcomes that were meta-analysed using the standardised mean difference approach (SMD), the MID was calculated by adding or subtracting 0.5 (given SD equals 1).

For the diagnostic questions, we assessed imprecision on the outcome of positive likelihood ratio because this was prioritised by the GDG as the most important diagnostic outcome for their decision making. The assessment of imprecision for the results on positive likelihood ratio followed the same concept as used in interventional reviews. For example, if the 95% confidence interval of the positive likelihood ratio crossed 2 zones (from moderately useful (5 to 10) to very useful (>10)) then imprecision was downgraded by 1, or if crossed 3 zones (not
useful (<5), moderately useful (5 to 10) and very useful (>10) then imprecision was
downgraded by 2.

3.3.3.7 Quality assessment of NMA

For the NMAs, quality was assessed by looking at risk of bias across the included evidence
(using standard GRADE approach for this domain), heterogeneity and inconsistency.

The following limits of the upper 95% CI for between-study standard deviation were used to
assess heterogeneity:

- less than 0.3 – low heterogeneity, quality of evidence not downgraded
- 0.3-0.6 – moderate heterogeneity, quality of evidence downgraded by one level
- 0.6-0.9 – high heterogeneity, quality of evidence downgraded by 2 levels
- 0.9-1.2 – very high heterogeneity, quality of evidence downgraded by 3 levels

Inconsistency in NMA has a different meaning than in pairwise meta-analysis, referring to the
discrepancy between direct and indirect evidence in closed treatment loops within the
network. If closed treatment loops existed then the following criteria were adopted:
- significant inconsistency in one loop – quality of evidence downgraded by one level
- significant inconsistency in more than 50% of loops where several loops exist – quality of
evidence downgraded by 2 levels

For fixed-effect NMAs that did not model heterogeneity, or for networks in which
inconsistency could not be assessed as no closed treatment loops existed, these criteria
were not considered to impact the quality of evidence.

3.3.3.8 Quality assessment of qualitative studies

Quality of qualitative studies (at study level) was assessed following the NICE checklists in
Methods Manual 2007. The main quality assessment domains are organised across the
definition of population included, the appropriateness of methods used and the completeness
of data analysis and the overall relevance of the study’ participants to the population of
interest for the guideline.

3.3.4 Use of absolute effect in decision making

The GDG assessed the evidence by outcome in order to determine if there was, or
potentially was, a clinically important benefit, a clinically important harm or no clinically
important difference between interventions. To facilitate this, binary outcomes were
converted into absolute risk differences (ARDs) using GRADEpro software: the median
control group risk across studies was used to calculate the ARD and its 95% CI from the
pooled risk ratio with the exception of estimation of baseline risk for breast cancer and CVD.

For breast cancer, baseline incidence for all women in the UK in 2010 was taken from the
Office of National Statistics (ONS) database. A limitation of using this statistic is that it
includes women on HRT in addition to those not on HRT. However, it was considered to be
the most reliable estimate available, as the proportion of women using HRT in the ONS
estimate is relatively low and the GDG indicated that the recording of prior HRT use in many
studies was unreliable. This annual incidence was then multiplied by 7.5, to reflect the
average length of follow-up in the studies included in the review, giving a baseline incidence
over 7.5 years of 9.45 per 1000 women. Breast cancer mortality was estimated similarly,
using 2011 data from the ONS database. The baseline incidence of mortality was estimated
to be 1.8 per 1000 women over 7.5 years.

For CVD there were a number of outcomes of interest for which it was necessary to estimate
baseline incidences. CHD incidence was obtained from a UK study by Weiner 2008 which
reported the rate in person-years of MI in women younger than 55 years and older than 55 years separately. A weighted average of these rates was calculated, and this was multiplied by the average length of included studies follow-up to give an incidence of CHD of 15 per 1000 people over 7.5 years. No information was found for the baseline incidence for the outcome of chronic heart disease (CHD) death. Therefore the incidence of CHD death and CHD were assumed to be equivalent, though results should be interpreted with caution due to unavailability of accurate baseline information for this outcome.

The rate of stroke was taken from the same UK study (Weiner 2008), and the incidence was calculated in the same way. The baseline incidence of stroke was 11.3 per 1000 women over 7.5 years. As the majority of strokes are ischaemic, the baseline incidence of ischaemic stroke was assumed to be the same. However, as haemorrhagic strokes are rarer, and UK data were not available for this outcome, we used the incidence in the control arm from any study reporting haemorrhagic stroke as the baseline risk.

As a composite of both MI and stroke, the incidence of CVD in untreated women was estimated to be the incidence of both stroke and MI, obtained by adding the rates from the Weiner 2008 study. This gave a baseline incidence of 26.3 per 1000 women over 7.5 years. The incidence of CVD death was considered to be equal to CVD.

As reliable UK data was not available for the incidence of fragility fractures, the incidence of CHD or CHD death in women with pre-existing heart disease, we used the incidence in the control arms from studies reporting this outcome in this population (default GRADE approach). The absolute risk therefore reflected the duration of the study/studies that contributed to these results and more information are provided as footnotes in the relevant tables.

3.3.5 Evidence statements

Evidence statements are summary statements that are presented after the GRADE profiles, summarising the key features of the clinical evidence presented. The wording of the evidence statements reflects the certainty or uncertainty in the estimate of effect. The evidence statements are presented by comparison (for interventional reviews) or by description of outcome where appropriate and encompass the following key features of the evidence:

- the number of studies and the number of participants for a particular outcome
- a brief description of the participants
- an indication of the direction of effect (if one treatment is beneficial or harmful compared with the other, or whether there is no difference between the 2 tested treatments)
- a description of the overall quality of evidence (GRADE overall quality).

3.3.6 Evidence of cost effectiveness

The aims of the health economic input to the guideline were to inform the GDG of potential economic issues related to diagnosis and management of menopause to ensure that recommendations represented a cost effective use of healthcare resources. Health economic evaluations aim to integrate data on benefits (ideally in terms of quality adjusted life years (QALYs), harms and costs of different care options.

The GDG prioritised a single review question, on managing the short-term symptoms of menopause, where it was thought that economic considerations would be particularly important in formulating recommendations and a review of the health economic literature was also undertaken for this question. For economic evaluations, no standard system of grading the quality of evidence exists and included papers were assessed using the economic evaluations checklist as specified in the guidelines manual. This literature review is presented in Appendix L and the evidence table is included in Appendix H.
Health economic reviews were undertaken for review questions relating to short-term treatment and symptoms, the diagnosis of premature ovarian insufficiency (POI) and the treatment of urogenital atrophy in women with menopause-related vaginal/urogenital atrophy.

No health economic literature review was reported for the long term risk and benefits of HRT. It was agreed that the economic analysis would not address the impact of HRT beyond 5 years because authors of studies considering a longer term impact have reported that cost-effectiveness is driven by differences in short-term symptom relief. In the context of this guideline it was not considered appropriate to investigate the cost-effectiveness of a treatment that looked only at a health economic evaluation of long term symptoms without considering the impact on short-term symptoms. Therefore relevant studies considering longer term risks and benefits would have been expected to have been captured by the systematic review we have planned on short-term treatments. However, the absence of a health economic review did not preclude the use of data from the clinical review of longer term risks and benefits in the health economic analysis if the GDG considered that there were important longer term risks and benefits from short term use of HRT.

No health economic review was undertaken on the review question focused on information and advice as this focused primarily on the content and quality of information that is routinely provided rather than whether the provision of information itself represent a cost-effective use of resources. Therefore, this question was not primarily about competing alternatives which have different opportunity costs and therefore was not considered suitable for a health economic review.

No clinical evidence was identified for classification systems for the diagnosis of menopause and it was thought a priori that it was most unlikely that there would be economic studies on this. Similarly, no clinical evidence was found on the intervals at which clinical review be undertaken to assess the effectiveness and safety of treatments to relieve menopausal symptoms and to determine when women need to be referred to specialist care and again it was thought a priori that it would be very unlikely that any economic evaluation would exist on this topic.

New economic analysis was undertaken by the health economist to address the cost effectiveness of HRT, non-HRT drugs, herbal preparations and other interventions given to women with vasomotor symptoms. This analysis is summarised in Section 7.6 and reported in full in Appendix L.

### 3.4 Developing recommendations

Over the course of the guideline development process, the GDG was presented with:

- evidence tables of the clinical and economic evidence reviewed from the literature. All evidence tables are in Appendix H
- summary of clinical and economic evidence and quality assessment (as presented in chapters 4 to 11)
- forest plots (Appendix J) and
- a description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (Appendix L).

Recommendations were drafted on the basis of the GDG interpretation of the available evidence, taking into account the balance of benefits, harms and costs between different courses of action. This was either done formally in an economic model, or informally. Firstly, the net benefit over harm (clinical effectiveness) was considered, focusing on the critical outcomes although most of the reviews in the guideline were outcome driven. When this was done informally, the GDG took into account the clinical benefits and harms when one intervention was compared with another. The assessment of net benefit was moderated by the importance placed on the outcomes (the GDG’s values and preferences), and the
confident the GDG had in the evidence (evidence quality). Secondly, it was assessed whether the net benefit justified any differences in costs.

When clinical and economic evidence was of poor quality, conflicting or absent, the GDG drafted recommendations based on their expert opinion. The considerations for making consensus-based recommendations include the balance between potential harms and benefits, the economic costs or implications compared with the economic benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues. The GDG also considered whether the uncertainty was sufficient to justify delaying making a recommendation to await further research, taking into account the potential harm of failing to make a clear recommendation.

The wording of recommendations was agreed by the GDG and focused on the following factors:

- the actions health professionals need to take
- the information readers need to know
- the strength of the recommendation (for example the word ‘offer’ was used for strong recommendations and ‘consider’ for weak recommendations)
- the involvement of patients (and their carers if needed) in decisions on treatment and care
- consistency with NICE’s standard advice on recommendations about drugs, waiting times and ineffective intervention

The main considerations specific to each recommendation are outlined in the ‘Recommendations and link to evidence’ sections within each chapter.

3.4.1 Research recommendations

When areas were identified for which good evidence was lacking, the GDG considered making recommendations for future research. Decisions about inclusion were based on factors such as:

- the importance to patients or the population
- national priorities
- potential impact on the NHS and future NICE guidance
- ethical and technical feasibility.

3.4.2 Validation process

This guidance is subject to a 6-week public consultation and feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders are responded to in turn and posted on the NICE website when the pre-publication check of the full guideline occurs.

3.4.3 Updating the guideline

Following publication, and in accordance with the NICE guidelines manual, NICE will undertake a review of whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update.

3.4.4 Disclaimer

Health care providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the
recommendations cited here must be made by practitioners in light of individual patient circumstances, the wishes of the patient, clinical expertise and resources.

The National Collaborating Centre for Women and Children’s Health disclaims any responsibility for damages arising out of the use or non-use of these guidelines and the literature used in support of these guidelines.

### 3.4.5 Funding

The National Collaborating Centre for Women and Children’s Health was commissioned by the National Institute for Health and Care Excellence to undertake the work on this guideline.
4 Diagnosis of perimenopause and menopause

4.1 Review question
What is the diagnostic accuracy of the following indicators (clinical and biological manifestations) in the diagnosis of perimenopause and menopause: menopausal symptoms (especially vasomotor), endocrine changes (specifically follicle-stimulating hormone, anti-Müllerian hormone, oestrogen or inhibin B) and total antral follicle count?

4.2 Introduction to topic
The most commonly used clinical definition of the phases of the climacteric considers:

- premenopause as menstrual cycling that is relatively normal for the women (bearing in mind that there is some gradual change in experience of menstruation across the life cycle for women e.g. alteration in cycle length, changes in period pain or premenstrual symptoms)
- perimenopause, also called the menopausal transition, is the interval in which many woman have irregular menstrual cycles before the menopause.
- a woman is defined as postmenopausal from one year after her last period. Within the UK population, the mean age of women who have a natural menopause is 51 years, although there is wide variation between women.
- menopause refers specifically to the last menstrual period but is rarely used as a diagnosis in itself because it is impossible to know if the menstruation is the last one; therefore postmenopause tends to be used more than menopause.

Terms such as climacteric, time of life or menopausal (as a general term) are probably less helpful as they are too broad to have clinical usefulness. Menopause if used in this section refers to the last menstrual period.

Current practice in the UK is to diagnose menopause clinically, on the basis of menstrual history and age. However, a number of other methods have been suggested as possible adjuncts or alternatives to a clinical diagnosis.

These definitions derive from the early World Health Organisation definitions (WHO 1994), which were elaborated for international use by the International Menopause Society (Utian 1999). Subsequently USA study teams (Gracia 2005; Harlow 2012) developed a more detailed staging of menopause, referred to as The Stages of Reproductive Aging Workshop (STRAW).

STRAW classification include additional criteria for defining specific stages of reproductive life. The revised staging system aims to provide a more comprehensive basis for classification and assessment, from the late reproductive stage through the menopausal transition and into postmenopause and this classification was the focus of another review question.

4.3 Clinical introduction
The aim of this question was to determine the diagnostic accuracy of age, menopausal symptoms, biochemical measurements (FSH, AMH, AFC, Inhibin B, Inhibin A, oestrogen) and ultrasound features (ovarian volume) to diagnose perimenopause and postmenopause. These indexes were considered either individually or in combination.
The diagnostic accuracy of these variables to diagnose the menopause is highly dependent upon the background population of women in whom the test has been applied. Therefore the evidence for the different tests investigated is presented against the background population of women.

For full details see review protocol in Appendix D.

4.4 Description of included studies


All included studies except 2 (Cooper 1995, Johnson 2004) defined menopause as being when amenorrhoea lasted for twelve or more months. Cooper 1995 used the definition of menopause as when FSH levels were elevated more than 15 IU/L whereas Johnson 2004 used a consensus based method using cycle irregularity and levels of serum FSH, LH, oestradiol, oestrone and progesterone.

The definition of premenopause was consistent across the included studies whereas perimenopause women were classified under different criteria (mainly related to differences in frequency of menstrual cycles).

4.4.1 Studies looking at the diagnostic accuracy of age and menopausal symptoms


A standardised questionnaire (such as the Menopause Rating Scale) was used to assess the prevalence of specific menopausal symptoms and their role in the diagnosis of menopause.

4.4.2 Studies looking at the diagnostic accuracy of biochemical measures

Four studies were included in this section.

Burger 1998 investigated the levels of inhibin A and B in premenopausal, perimenopausal and postmenopausal women aged 48 to 59 years. This study was conducted in Australia and included a subset of 110 women from a larger study (The Melbourne Women’s Mid-Life Health Project).

Two studies (Stellato 1998, Henrich 2006) looked at the role of FSH in the diagnosis of menopause. Henrich and colleagues assessed the usefulness of FSH in determining menopausal status in women participating in the National Health and Nutrition Examination Survey (NHANES). Both studies were conducted in the USA and included premenopausal, perimenopausal and postmenopausal women.
Shin 2008 assessed the usefulness of a variety of hormonal markers (oestradiol, FSH, AMH and inhibin) to determine menopausal status for 144 postmenopausal women aged 50-59 years and premenopausal women aged 20 to 49 years.

4.4.3 Studies looking at the diagnostic accuracy of ultrasound features

Giacobbe 2004 assessed the usefulness of age and ovarian ultrasonography, which measured antral follicle count and ovarian volume to diagnose menopausal status. The study included women aged 40 to 55 years old in Brazil. Only 2 groups of women were included – postmenopausal women and women who were not yet menopausal (described as premenopausal but actually including any women who had not had 12 months amenorrhoea, for example both premenopausal and perimenopausal women).

4.4.4 Studies looking at combination tools

One study (Johnson 2004) tested the usefulness of 3 different algorithms for the diagnosis of perimenopause and menopause in 507 women aged 21 and 55 years who were under investigation for suspected myocardial ischaemia. Two of these algorithms were previously developed, a menstrual algorithm (based on menstrual history alone) and a historical algorithm (based on menstrual history, surgical history and age). The third algorithm (hormonal algorithm) was developed as part of this study and was based on menstrual history, surgical history, age and hormone levels (FSH and oestradiol). Premenopausal, perimenopausal and postmenopausal women participated in this study.

4.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F., study evidence tables in Appendix H.; forest plots in Appendix J.; and exclusion list in Appendix G.

The accuracy of the different diagnostic tests is dependent on the population of women in whom the test is conducted. For example, the specificity of a given test to distinguish postmenopausal women from a population of perimenopausal and postmenopausal women will be different to the specificity of the test when conducted in a population which also includes premenopausal women. This also changes the positive and negative likelihood ratios. Therefore, separate GRADE tables are presented to reflect the evidence for distinguishing perimenopausal and postmenopausal women from different background populations.

The type of menopause symptoms along their duration was reported differently across the studies. For example, vasomotor symptoms were reported as hot flushes, cold sweats, night sweats, palpitations or a combination of these symptoms. Meta-analysis was not conducted and results are reported separately by symptom due to differences in reporting.

Likelihood ratios are reported as the primary measure of diagnostic accuracy. The positive likelihood ratio reports the number of times more likely postmenopausal women are to have that symptom than other women (either premenopausal women, perimenopausal women or both). The higher the value, the more likely it is that a woman with a positive test is menopausal. By convention, a value between 5 and 10 is regarded as moderately useful and a value of 10 and over is very useful. Tests where the likelihood ratios lie close to 1 have little practical significance.

The negative likelihood ratio indicates whether the absence of a sign, age band, or endocrine level is a good way of determining that a woman is not in the menopause. The lower the value, the more likely it is that a woman with a negative test is not menopausal. In this case, the lower the value reported in the GRADE table the better the test may be to diagnose menopause by ruling out cases that are not menopausal. By convention, a value of < 0.1 is
regarded as very useful, and a value of 0.1 to 0.2 is moderately useful. Again, a negative
likelihood ratio close to one demonstrates that a negative test is equally likely in both
menopausal and non-menopausal women.

The moderately useful and very useful results are reported in the summary GRADE tables
below (white for moderately useful, grey for very useful).

A summary of the findings is also presented in the following graphs for easier interpretation
separately for single and combination tests (grey colour demonstrates useful test, dark not
useful and white neutral).

**Figure 3: Single tests for diagnosis of menopause (background population:
perimenopause) - results on positive likelihood ratio**
Figure 4: Single tests for diagnosis of menopause (background population: perimenopause) - results on negative likelihood ratio

Figure 5: Single tests for diagnosis of menopause (background population: premenopause) - results on positive likelihood ratio
Figure 6: Single tests for diagnosis of menopause (background population: premenopause) - results on negative likelihood ratio

![Negative Likelihood Ratio Graph](image)

Figure 7: Single tests for diagnosis of menopause (background population: premenopause plus perimenopause women) - results on positive likelihood ratio

![Positive Likelihood Ratio Graph](image)
Figure 8: Single tests for diagnosis of menopause (background population: premenopause plus perimenopause women) - results on negative likelihood ratio

4.6 Economic evidence

No health economic search was undertaken for this question. The GDG’s prior view was that these tests are often performed unnecessarily and that the topic was included in the scope as a potential area for disinvestment.

4.7 Evidence statements

Background population perimenopausal women

Low quality evidence from one study found that if a woman was aged 55 years or more then she is more likely to be menopausal but being aged less than 55 did not reduce the chances of being menopausal as well. It is not useful to distinguish menopause from perimenopause only based on the age criterion if this is less or equal than 45 or 50 years (moderate to low quality evidence).

Moderate to very low quality evidence from several studies which reported vasomotor symptoms (presenting as hot flashes or nights sweats) in different time points (the last 2, 4 weeks or 12 months) or without a specified time point concluded that the presence of vasomotor symptoms was not found a useful tool to distinguish menopause from perimenopause. One study reported that having a detectable level of inhibin A reduced the chance of being postmenopausal, while having an undetectable level of inhibin A did not increase the chance of being menopausal (moderate quality evidence). Moderate quality evidence for other studies found that no other endocrine tests (FSH or inhibin B) was useful to distinguish menopause from perimenopause.
Background population premenopausal women

Moderate quality evidence from one study reported that if a woman was aged 45 years or more this criterion had no impact on the chances of being peri or menopausal but being aged less than 45 reduced the chances of being menopausal. On the contrary, if a woman was aged 50 to 55 years or more she was more likely to be peri or menopausal and being aged less than 50 reduced the chances of being menopausal (moderate quality evidence). Finally, one study reported that if a woman was aged 60 or more then she was more likely to be menopausal and being aged less than 60 did not reduce the chances of being menopausal.

Moderate to low quality evidence from 2 studies reported that hot flushes and night sweating (over an unspecified time period) increased the chances of being peri or menopausal but having none of these symptoms did not reduce the chances of being menopausal. However, low quality evidence from another 2 studies reported that hot flushes and night sweating at different time points (the last month, or over an unspecified time period) was not useful to distinguish peri or menopause from premenopausal women.

A meta-analysis of 2 studies reported that current hot flushes or night sweats did not increase the chances of being menopausal but having no current hot flushes reduced the chances of being menopausal. The evidence for this finding was of very low quality. The presence of other vasomotor symptoms was not useful to distinguish menopause from premenopause.

Low quality evidence from one study showed that rapid heart beating (palpitations) (over an unspecified time period) might increase the chance of being menopausal.

Moderate to very low quality evidence from different studies looking at the diagnostic accuracy of biochemical measures to diagnose menopause found that:

- FSH level of > 22.3mIU/mL increased the chances of being menopausal while a level < 22.3mIU/mL reduced the chances of being menopausal (low quality evidence)
- AMH level of < 0.5ng/mL increased the chances of being menopausal while a level > 0.5ng/mL reduced the chances of being menopausal (low quality evidence)
- oestradiol level of < 34.5pg/mL (equivalent to 126.6 pmol/mL) increased the chances of being menopausal while a level >34.5pg/mL reduced the chances of being menopausal (very low quality).
- detectable level of inhibin A reduced the chance of being menopausal while having an undetectable level of inhibin A did not increase the chance of being menopausal (moderate quality evidence).
- inhibin B level of <0.4pg/mL increased the chances of being menopausal while a level > 0.4pg/mL reduced the chances of being menopausal. The evidence for this finding was of low quality.

Background population all women

Moderate quality evidence reported that if a woman was aged 45 years or more then this had no impact on the chances of being postmenopausal but being aged less than 45 reduced the chances of being menopausal. A meta-analysis of 2 studies reported that if a woman was aged 50 years or more then she was more likely to be postmenopausal but being aged less than 50 did not reduce the chances of being menopausal (very low quality evidence). The same conclusion came from another study which looked at the cut-off points of 55 and 60 years (both were of moderate quality).

A pooled analysis of 2 studies found that current hot flushes or night sweats did not increase the chances of being menopausal but the absence of hot flushes or night sweats reduced the chances of being menopausal. The evidence for this finding was of low quality. The presence
of other vasomotor symptoms was not useful in distinguishing postmenopausal women from all other women (moderate to very low quality evidence).

Moderate quality evidence from one study found that having a detectable level of inhibin A reduced the chance of being menopausal while having an undetectable level of inhibin A did not increase the chance of being menopausal, whereas inhibin B was found to be not useful to diagnosis menopause.

No ovarian ultrasound features (antral follicle count ≤ 2 follicles or ovarian volume < 4cm³) were found useful to distinguish menopausal women from all other women (low quality evidence).

**Combinations of variables or algorithms**

Low to very low quality evidence from one study found that all algorithms either menstrual (classifying women according to the time since their last period – either within 3 months, within 3 to 12 months, or longer than 12 months ago), hormonal (classifying women according to their menstrual history, surgical history, age, FSH and oestradiol levels) or historical (classifying women according to their menstrual history, surgical history and age) allowed for the correct classification of peri or postmenopausal women.

The following conclusions from single studies were made regarding the usefulness of tools for diagnosis of perimenopausal women from postmenopausal women:

- being aged less than 55 or 60 years may reduce the chances of being perimenopausal but being over 55 did not increase the chance of being perimenopausal (very low quality evidence). No other age groups (< 45 years or < 50 years) were useful to distinguish perimenopausal from postmenopausal women.
- the presence of vasomotor symptoms alone was not useful to distinguish perimenopausal from postmenopausal women (moderate to very low quality evidence).
- no endocrine tests (inhibin A or inhibin B) were found useful to distinguish perimenopausal from postmenopausal women (moderate quality evidence).

The following conclusions from single studies were made regarding the usefulness of tools to diagnosis of perimenopausal women from premenopausal women:

- a woman aged 45 years or more may not have an increased chance of being perimenopausal, but being aged less than 45 reduced the chances of being perimenopausal (moderate quality evidence). The same study also showed that being a women aged 55 years or more had an increased chance of being perimenopausal but being aged less than 55 did not reduce the chance of being perimenopausal. No other age groups (≥ 42 years, ≥ 46 years, ≥ 50 years, ≥ 60 years) were found to be useful to distinguish perimenopausal women from premenopausal women (moderate to very low quality evidence).
- one or more hot flushes or night sweats per day during the past 6 months may increase the chances of being perimenopausal while the absence of hot flushes or night sweats did not reduce the chances of being perimenopausal (very low quality). The presence of other vasomotor symptoms alone was not found to be useful to distinguish perimenopausal women from premenopausal women (moderate to very low quality evidence).
- FSH level of >13 mIU/mL increased the chances of being perimenopausal but a level < 13 mIU/mL did not reduce the chances of being perimenopausal. The evidence for this finding was of low quality. No other endocrine tests (FSH level of ≥24 IU/L, inhibin A or inhibin B, AMH, oestradiol) were found to be useful to distinguish perimenopausal women from premenopausal women (moderate to low quality evidence).
- the presence of at least one of a list of symptoms increased the chances of a woman being perimenopausal. The following symptoms were included: at least one hot flush or night sweat per day for the past 6 months. However, not reporting any of these symptoms
did not reduce the chances of being perimenopausal. The evidence for this finding was of moderate quality. No other combination tests were found to be useful to distinguish perimenopausal women from premenopausal women.

Moderate to very low quality evidence did not find either the presence of vasomotor symptoms or endocrine tests useful tools to distinguish perimenopausal women from all other women whereas other evidence of moderate quality showed that both a menstrual algorithm (classifying women according to the time since their last period – either within 3 months, within 3 to 12 months, or longer than 12 months ago) and a hormonal algorithm (classifying women according to their menstrual history, surgical history, age, FSH and oestradiol levels) allowed for the correct classification of perimenopausal women.

### 4.8 Evidence to recommendations

#### 4.8.1 Relative value placed on the outcomes considered

The GDG has considered all the properties of diagnostic accuracy measurements for decision making in this topic: sensitivity, specificity, positive and negative likelihood ratio and Area under the Curve (AUC). The GDG considered the relative importance of having a high false positive and high false negative result from the diagnosis of menopause and consequences in women’s further clinical management.

Likelihood ratios were considered the most critical measures of diagnostic accuracy of different tests for menopause and for the GDG’s decision making. The positive likelihood ratio reports the number of times more likely postmenopausal women are to have that symptom than non-menopausal women (either premenopausal women, perimenopausal women or both, depending on the study). The higher the value, the more likely it is that a woman with a positive test is postmenopausal.

Given that women at different stages of menopause (peri or postmenopause) may experience different symptoms and require different further management it was considered important by the GDG to examine the role of each test to diagnose different stages of menopause in reference to the background population.

#### 4.8.2 Consideration of clinical benefits and harms

Different tests were considered to diagnose peri or postmenopausal women from different background populations (premenopause, all women). In summary no indication (age, vasomotor symptoms, biochemical measures, endocrine changes, ultrasound features measuring ovarian volume) when they were examined in isolation were found to accurately discriminate those women who have positive and negative diagnosis. This was the case when different background populations were taken into consideration. Some indicators such as age above 55 or 60 years were found to have useful positive likelihood ratio but not very useful negative likelihood ratio.

On the other hand, algorithms either as combinations of menstrual (classifying women according to the time since their last period – either within 3 months, within 3 to 12 months, or longer than 12 months ago), hormonal (classifying women according to their menstrual history, surgical history, age, FSH and oestradiol levels) and historical (classifying women according to their menstrual history, surgical history and age) allowed for the correct classification of menopausal women from both premenopausal and perimenopausal women.

The GDG discussed the role of hot flushes in diagnosis of menopause as these are considered to be one of the principal symptoms for visiting a health care professional for women around the age of menopause in the UK. The GDG were surprised that it did not produce high diagnostic accuracy for menopause as a single measurement. They considered that this may be because this symptom also occurs in a significant number of premenopausal
Menopause
Diagnosis of perimenopause and menopause

1. women, so hot flushes are considered together with other indications such as absent or infrequent menses to accurately distinguish menopause from premenopause.

2. The reviewed evidence did not give the group confidence to decide that the diagnosis of menopause should involve the use of biochemical, hormonal tests or an ultrasound test for assessing the function of the uterus as the results did not provide robust evidence for their routine use in diagnosis of menopause. The GDG concluded that the age (over 45 years) with the combination of amenorrhea for at least 12 months should be considered adequate for diagnosis of post menopause and oligomenorrhea for the diagnosis of perimenopause. This is currently the routine clinical method of diagnosis of menopause.

3. The GDG also considered that FSH measurement may occasionally be useful for the diagnosis of menopause and perimenopause for those cases of women over the age of 45 years who do not have typical menopause symptoms such as vasomotor but have atypical symptoms of recent onset. However they highlighted that there is a lack of precision of FSH measurements in the perimenopause as it fluctuates considerably over short periods of time during the years leading up to the menopause.

4. The diagnostic accuracy of FSH as a tool for menopause may be confounded for those taking hormonal treatment and the group decided to inform prescribers that FSH levels should not be considered for measurement for this group of women. Many women will experience irregular bleeding or absent menstruation when using hormonal contraception. Some, for example those on injectable progestogens, may also experience menopausal symptoms since they are hypoestrogenic. The group concluded that there is no value in measuring gonadotrophins (LH, FSH) in these women since they will be altered by the hormonal contraceptive. In addition, women on HRT will have decreased gonadotrophin levels.

4.8.3 Consideration of economic benefits and harms

1. Diagnosis carries an opportunity cost, with the resources used for it unavailable for alternative use within the NHS. Therefore, it is important that diagnosis ultimately leads to improved management and outcomes. However, the GDG were not persuaded by the clinical evidence alone that there was a place for the routine use of biochemical, hormonal tests or ultrasound test for uterus function in diagnosis of menopause. Therefore, it is reasonable to conclude that such tests do not represent an efficient use of scarce NHS resources.

4.8.4 Quality of evidence

1. The majority of evidence contributed to this section was moderate to low as assessed by the QUADAS-2 checklist. The thresholds of measurements were not a priori selected based on clinical considerations but the results reported as per studies. The studies varied considerably in terms of population characteristics but this is not unusual for diagnostic studies.

4.8.5 Other considerations

1. The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

2. The GDG has discussed the point that the diagnosis of perimenopause would be the only clinically relevant diagnosis for implications on further management among women presenting with any type of menopausal symptoms.

4.8.6 Key conclusions

1. The GDG concluded that:
• diagnosis of the perimenopause would be the only useful diagnosis clinicians should consider making

• age and amenorrhea are sufficient clinical indicators for the routine diagnosis of menopause

• biochemical measurements, hormonal tests and ultrasound tests were not found useful in routine practice of diagnosis of menopause and perimenopause.

4.9 Recommendations

1. Diagnose the following without laboratory tests in otherwise healthy women aged over 45 years with menopausal symptoms:
   • perimenopause based on vasomotor symptoms and irregular periods
   • menopause in women who have not had a period for at least 12 months
   • menopause based on symptoms in women without a uterus

2. Take into account that it can be difficult to diagnose menopause in women taking sex steroids.

3. Do not use the following laboratory and imaging tests to diagnose perimenopause or menopause in women aged over 45 years:
   • anti-Müllerian hormone
   • inhibin A
   • inhibin B
   • oestradiol
   • antral follicle count
   • ovarian volume.

4. Do not use a serum follicle stimulating hormone (FSH) test to diagnose menopause in women using combined oestrogen and progestogen contraception or high-dose progestogen.

5. Consider using a FSH test to diagnose menopause only:
   • in women aged over 45 years with atypical symptoms
   • in women aged 40 to 45 years with menopausal symptoms, including a change in their menstrual cycle
   • in women aged under 40 years in whom menopause is suspected (see also section 11.1).
5 Classification systems for the diagnosis of menopause

5.1 Review question

What is the usefulness of formal classification systems compared with non-structured classification systems in the diagnosis of menopause and in guiding further treatment?

5.2 Introduction to topic

There are a variety of classification systems to diagnose the menopause and perimenopausal symptoms. These have been used as research tools. In Primary Care the diagnosis of the menopause in a woman over 45 years of age is still principally done by taking a good medical history, listening to the woman’s symptomatic complaints, and excluding other possible diagnoses where appropriate. The GDG looked at the evidence of how useful these various research tools of classification systems might be to the general clinician, and to the woman herself to help with treatment options.

5.3 Clinical introduction

A number of classification systems have been developed to define stages of the menopausal transition, (such as STRAW (Stages of Reproductive Aging Workshop), STRAW 10 and RESTAGE Algorithm), largely as an aid to research. However, the focus of this question was to assess whether these classification systems are also of use in clinical practice. The aim of this review was to identify whether the use of structured classification systems are useful tools to assess different stages of the menopause by guiding further investigation and treatment for menopausal symptoms, over using a clinical history alone.

The outcomes prioritised by the GDG were: correct diagnosis of menopause, guidance for further investigation or treatment and health related quality of life (HRQoL).

For full details see review protocol in Appendix D.

5.4 Description of included studies

The search for this topic included both RCTs and comparative cohort studies, but no studies were identified which met the inclusion criteria.

5.5 Clinical evidence profile

No evidence profile was generated.

5.6 Economic evidence

No health economic studies were identified and no health economic modelling was planned for this question.

5.7 Evidence statements

No studies were identified for this review question and therefore there is no evidence profile.
5.8 **Evidence to recommendations**

5.8.1 **Relative value placed on the outcomes considered**

The outcomes prioritised by the GDG for this review question were the correct diagnosis of menopause, guidance for further investigation or treatment and HRQoL.

5.8.2 **Consideration of clinical benefits and harms**

Given the absence of clinical evidence for this review topic and the lack of use of any of these classification systems in current routine clinical practice, the GDG agreed that no recommendation could be made within this section.

The GDG discussed that they used standard questions to classify women in the different phases of menopause such as time since last period and age but did not apply a formal classification system. The group also discussed some of the limitations of these classification systems, for example STRAW which was developed to apply only to healthy women and cannot apply to some common groups of women in menopause such as women who had undergone hysterectomy (as some of the criteria used in the tools were bleeding criteria) or those with a high BMI. Therefore, the GDG did not consider making a recommendation in favour or against use of these classification tools.

5.8.3 **Consideration of economic benefits and harms**

This review aimed to compare different classification systems and did not find any clinical evidence to meet this protocol. Furthermore, none of these classification systems are routinely used in clinical practice. Any classification system would impose some opportunity cost through its administration and therefore in the absence of any evidence of benefit they cannot currently be considered to represent value for money in the NHS in routine clinical practice.

5.8.4 **Quality of evidence**

No clinical evidence was found for this review question.

5.8.5 **Other considerations**

None of the GDG members use a classification system in routine practice. GDG members used standard questions – such as time since last period, or age etc. However, a classification system could be useful in women with POI as they often experience delays in diagnosis and treatment.

5.8.6 **Key conclusions**

The GDG concluded that the absence of evidence and the GDG's lack of clinical experience in the use of these classification systems led to not making recommendations in this topic.
6 Information and advice

6.1 Review question

What are the information needs for women in menopause?

6.2 Introduction to topic

The menopause is a natural milestone in women’s lives, and can be seen as the gateway to further aging processes.

Most women are aware of the possibility of hot flushes and night sweats around menopause, but most are unaware of the increased risk of a number of health conditions, e.g. heart disease, osteoporosis, urinary incontinence, vaginal atrophy and decreased sexual function. It is important that information is available to women so that they can make informed choices about their lifestyle and potential treatments.

There are many different options for women, including lifestyle changes, complementary medicines, prescribed medicines or doing nothing and letting time pass, to reduce the symptoms. Every action or inaction has benefits and negative sequela associated with it. It is important that the woman understands the consequences of her decisions, and is able to make an evidence-based choice that she is comfortable with. Whatever her decision, it needs to be noted in her primary care records, if a discussion has taken place, and the information provided. The challenge for health professionals is to provide this highly complex data in a way tailored for that individual. Every woman will have a different view of themselves, and what is important to them. That individual woman’s information needs also change as time goes by, so the situation has to be reassessed each time she requests new information, or if new symptoms develop, or disappear.

6.3 Clinical Introduction

The aim of this review was to establish the most common areas of information needs for women in menopause and what the most effective ways of delivering these information are. The focus population of this review question was peri and postmenopausal women. Information was presented separately for the following subgroups if data were available:

- women with POI
- women with iatrogenic menopause, particularly due to cancer treatment or those at risk of cancer
- women with natural menopause who present for symptom-relief

For the first part of the question, systematic reviews of qualitative studies, observational studies (ideally large cohorts), and qualitative studies (natural history data, patient reported outcomes) were considered for inclusion. Areas of information needs were the focus of this part of the review question.

For the second part of the question, both randomised controlled trials (RCTs) and comparative cohort studies were selected for inclusion. Qualitative studies could also provide supplementary information for this part of the review question. Any format of delivery of information including written, oral communication, and websites regarding menopause was considered for the second part of this review question. Patient knowledge and number of visits to health care professionals were selected as the outcomes for this part of the review question. For full details see review protocol in Appendix D.
6.4 Description of included studies

A total of 28 studies were included in this review, some of which overlapped in both parts of the review question:

1st part: areas of information needs of women in menopause:

- 12 studies assessed different areas of information needs for women in menopause (Armitage 2007; Alfred 2006; Connelly 1999; Thewes 2003; Hallowell 2000; Fox-Young 1995; Mahon 2000; Mingo 2000; Theroux 2007; Walter 2004; Wathen 2006; Roberts 2001).

2nd part: effectiveness of different information delivery methods:

- 20 studies examined the effectiveness of different methods of information provision; 10 of these were RCTs (Becker 2009; Deschamps 2004; Legare 2008; Murray 2001; Rothert 1997; Kiatpongsan 2014; Rostom 2002; Forouhari 2010, Liao 1998; Hunter 2009), 1 was a comparative observational study (Fortin 2001) reporting results from survey questionnaires; and 9 studies had qualitative elements (Andrist 1998; Bravata 2002; Clinkingbeard 1999; Legare 2007; Roberts 1991; Wathen 2006; Doubova 2012; Walter 2002; Thewes 2003) and were used as supplementary information.

The methods of information provision in the included studies varied from booklets to educational courses. The effectiveness of these methods was assessed by the RCTs using a decision conflict score (Becker 2009; Deschamps 2004; Legare 2008; Murr ray 2001; Rothert 1997), a knowledge score (Becker 2009; Legare 2008; Kiatpongsan 2014; Rostom 2002) or a quality of life score (Forouhari 2010).

Studies were conducted in the following countries: USA (11), Canada (5), UK (7), Australia (3), Iran (1), and France (1).

The majority of women in the included studies were in natural menopause with the exception of Mahon 2000 which included women undergoing early menopause due to cancer treatment, and Hallowell 2000 which recruited women in surgical menopause.

Full details of the included studies are given in Appendix F. A summary of the main information needs areas that are covered in the qualitative studies is presented in Table 10.

Table 10: Studies in which information on the following was found helpful by women – or would have been helpful if they had received it

<table>
<thead>
<tr>
<th>STUDIES</th>
<th>Diagnosis of menopause</th>
<th>Menopausal symptoms</th>
<th>HRT (benefits, risks, optimum length of treatment, withdrawal options)</th>
<th>Unbiased explanation of guidelines and latest research</th>
<th>Self-management strategies including lifestyle changes</th>
<th>Non HRT options</th>
<th>How family history affects risk</th>
<th>Source of reliable information</th>
<th>Fertility issues</th>
<th>Sexuality</th>
<th>Sources of emotional support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfred 2006</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Andrist 1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armitage 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bravata 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinkingbeard 1999</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connely 1999</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 6.6 Economic evidence

No health economic search was undertaken for this review question and consequently no evidence was found. This question focused on the content and quality of information that is routinely provided rather than whether the provision of information itself represent a cost-effective use of resources. Therefore, this question is not primarily about competing alternatives which have different opportunity costs and therefore was not considered suitable for a health economic review.
6.7 Evidence statements

Areas of information needs

Moderate to very low quality evidence from the 12 included qualitative studies (employing interview or survey designs) showed that the areas of information needs identified by women in menopause were consistent. Anxiety over hot flushes was rarely reported as an information need, except in the case of one study in which women in iatrogenic menopause felt that there was lack of information regarding the use of HRT before surgery (moderate quality evidence). Low quality evidence from one study found fertility was an important issue for younger women, and they needed to discuss this with their HCPs. This was especially the case for women with iatrogenic menopause as their doctors gave fertility a low priority when treating diseases such as cancer.

The most widely reported area of information needs was related to the use of HRT. Moderate to very low quality evidence from six studies showed that women would like more guidance and evidence around HRT provided by their health professionals, especially because the available information on the internet is confusing regarding the benefits and risks. Although low quality evidence from one study showed that women expressed their preference that their health professionals should be more involved in the decision making on whether to use HRT some other women felt that a presentation of “facts” around the benefits and side effects of HRT would be a more appropriate way of information provision.

In terms of who is the most suitable information provider, moderate to low quality evidence from two studies showed that women noted that health care professionals should provide this information regarding menopause. Moderate to low quality evidence from another two studies reported that women felt that specialist doctors may be more helpful than GPs whereas low quality evidence from another study reported that women expressed that doctors were too busy to see them (as they were considered not ‘ill’). Moderate to low quality evidence from two studies found that a peer support system (in a physical group or through reported testimonials) could be another effective method of communication.

Methods of information provision

Booklet (tailored decision aid booklet with information on risk factors of diseases, current guidelines; symptoms of menopause, treatment options including HRT)

Moderate quality evidence from three individual RCTs with more than 300 participants showed no significant difference in decision conflict scores as a measure of personal perceptions of effective decision making for the use of HRT among women who have used the booklet compared with those who haven’t used it. Two RCTs also found no significant difference in the knowledge scores of HRT risks between the two comparison groups.

Enhanced booklets (interactive multimedia programme & booklet; DVD & booklet)

Moderate quality evidence from one RCT with more than 400 participants showed a significantly higher knowledge score in HRT risks in women who have used a booklet compared with those who haven’t used it. Another RCT found that the impact on women’s decision making does not seem to be significantly higher with the use of enhanced booklets (moderate quality evidence).

Information provider

Low quality evidence from one RCT with over a hundred women showed that there was no significant difference in the decision making experience about HRT between those who
received information from a pharmacist compared to those receiving information from a booklet.

**Educational courses**

Low to very low quality evidence from three 4 individual RCTs found that women attending an educational course about menopause had a positive impact on their decision making (as assessed with a higher decision conflict score) and lowering the level of uncertainty about menopause (increasing knowledge) compared to those who were given a booklet or not attending a course.

**Healthcare professional consultations (supplementary qualitative information)**

Doctors were seen as a useful source of information, however 3 studies found they sometimes lacked sympathy, or had strong opinions, or were not understood due to short consultations and verbal-only communication. Two qualitative studies also showed that women were also keen to self-manage, especially in a peer-group setting. Other studies found that specialist doctors were more helpful than GPs. Overall the quality of this evidence was of low to very low quality.

**Peer information provision (supplementary qualitative information)**

Two qualitative studies found that elements of peer work (in a physical group or through reported testimonials) were an effective method of communication (moderate to very low quality evidence).

**Risk presentation (supplementary qualitative information)**

One qualitative study indicated women prefer bar chart presentations to other formats as part of graphical presentation of risk than textual presentation of risk for the effects of HRT. Women also reported their preference on life-time survival information about the HRT risks. Two further studies emphasised women's need for clear factual information with which to assess risk for themselves.

6.8 **Evidence to recommendations**

6.8.1 **Relative value placed on the outcomes considered**

The main outcome for the first part of the question was the exploration of areas of information needs for women in menopause. For the second part of the question, the GDG considered patient knowledge and number of visits to health care professionals as the most important outcomes to assess the effectiveness of different types of information provision.

6.8.2 **Consideration of clinical benefits and harms**

The included evidence showed the importance of clear information provision regarding the diagnosis of menopause and its related associated symptoms. Evidence showed that some women often did not feel comfortable raising the topic of menopause with clinicians and wanted the clinician to raise it with them instead. Hormone replacement therapy and its associated benefits and risks was widely reported in the studies as a common theme of information needs for women in menopause but also some women wanted to know more information about the non HRT options for relief of menopausal symptoms. Fertility, sexuality and finding sources of emotional support were other areas of information needs identified by the women who participated in the included studies.
Women in the included studies also noted that specialist health professionals may be more helpful in provision of information than general health providers but the results were not consistent. Some evidence also showed that peer support groups can be another useful source of information for menopause. Sources of reliable evidence are particularly important in the area of the menopause and HRT as this continues to be a popular area with the press and information provided not is always completely accurate.

The GDG discussed the findings of these studies and decided that women in menopause should be given specific information about the different stages of menopause, the most common symptoms they may experience, how menopause is diagnosed, and the associated benefits and risks of available treatments. The GDG also discussed that, when women are contacting healthcare professionals regarding menopause, this may be an opportunistic window for healthcare professionals to highlighting the importance of lifestyle changes and promote good health and wellbeing (such as general screening of blood pressure and lipids as well as they attend for mammograms and smear tests as part of the national screening programmes).

In relation to menopausal symptoms, the GDG wished to elaborate on the most common menopausal symptoms women may experience including the change in their menstrual cycle, vasomotor and, musculoskeletal symptoms, mood disturbance, urogenital problems and sexual difficulties.

The GDG’s expert opinion was that information provision and support on the risk of impaired fertility and early menopause was important to woman undergoing medical treatment such as chemoradiation, women undergoing gynaecological surgery or woman at high risk of cancer. The GDG made a separate recommendation for this group of women, and considering referral to a healthcare professional with expertise in menopause owing to the complexity of their needs.

From the included evidence, enhanced booklets with digital presentation of information were considered useful to increase women’s knowledge about different aspects of menopause but they were not necessarily found helpful in their decision-making about different treatment options when compared to a group of women who did not have this type of information. However, booklets for the use of HRT were not found to be either useful in increasing the decision-making ability of women on selecting the most appropriate treatment for menopausal symptoms, or necessarily knowledge around the topic. Lectures and educational courses were also found helpful to increase women’s knowledge about menopause and its treatment options compared to other types of information provision.

The GDG discussed the different formats of presenting information, including using evidence based electronic sources, and they concluded that different styles of information may work differently for women from different socio-demographic backgrounds and health professionals should be flexible to use a variety of formats of information provision to best meet an individual woman’s needs.

### 6.8.3 Consideration of economic benefits and harms

Providing advice is a standard part of routine clinical practice. It typically involves a small opportunity cost in terms of staff time and consumables. There is a potential gain from avoiding discontinuation of HRT when women have a better understanding of side effects of HRT and unnecessary re-consultations. The Group discussed the value of peer support groups as information providers in the studies but did not identify strong enough evidence for an intervention which would have perhaps had more significant associated costs.
6.8.4 Quality of evidence

The quality of the evidence for the first part of the question, which mainly included qualitative studies, was low to very low quality although some studies employed appropriate methods of data analyses for these studies such a phenomenological or grounded theory approach. However the interpretation of results of this evidence is often restricted from the small sample size of the studies, the low response rate and the lack of generalising the results given the studies were conducted in specific population groups.

For the second part of the review question, randomised and observational studies were included. The quality evidence was moderate to very low as there was serious risk of bias involved and lack of data available in order to precisely estimate relative and absolute effects of the interventions (some studies only provided p values). In addition, these studies primary focus was often on other topics and they only included a small amount of detail about information provision.

6.8.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

Components of good patient experience in general are set out in ‘Patient experience in adult NHS services’. GDG were aware of this related NICE guidance and so focussed on recommendations specific to women undergoing menopause.

Information regarding the short and long term impact of HRT is given separately in this guidance (Chapters 7, 10–10.8).

6.8.6 Key conclusions

The GDG concluded that different areas of information provision are important for women in menopause including menopause symptoms, different treatment options and benefits and risks associated with the use of HRT. There may be specific information needs for women with POI and iatrogenic menopause such as information about fertility. Different presentations of information may be helpful as aids of decision tools for women in menopause.

6.9 Recommendations

6. Give information to menopausal women and their family members or carers (as appropriate) that includes:
   - an explanation of the stages of menopause
   - common symptoms (see recommendation 8) and diagnosis
   - lifestyle changes and interventions that could help general health and wellbeing
   - the benefits and risks of treatments for menopausal symptoms.

7. Give information on menopause in different ways to help encourage women to discuss their symptoms and needs.

8. Explain to women that as well as a change in their menstrual cycle they may experience a variety of symptoms associated with menopause, including:
   - vasomotor symptoms (for example, hot flushes and sweats)
   - musculoskeletal symptoms (for example, joint and muscle pain)
- effects on mood (for example, low mood)
- urogenital symptoms (for example, vaginal dryness)
- sexual difficulties (for example, low sexual desire).

9. Offer women who are likely to go through menopause as a result of medical or surgical treatment (including women with cancer, at high risk of hormone-sensitive cancer or having gynaecological surgery) support and:
   - information about menopause and fertility before they have their treatment
   - referral to a healthcare professional with expertise in menopause.
7 Managing short-term symptoms

7.1 Review question
What is the most clinical and cost-effective treatment for the relief of individual menopause-related symptoms for women at menopause?

7.2 Introduction to topic
Menopausal symptoms are extremely common. Hot flushes and night sweats are the most common symptoms in women living in the UK, and in addition many women report other symptoms which can include sleep disturbance, depression and mood changes, musculoskeletal pain, and urogenital symptoms. Sexual problems around the menopause include vaginal dryness and dyspareunia, and low libido; although these are complex symptoms, hormonal changes are often a contributing factor. It is less clear whether anxiety, irritability, palpitations, skin dryness, and fatigue can be attributed directly to the menopause; fatigue may be due to sleep disturbance from night sweats. The duration and severity of symptoms experienced are not uniform; symptoms may develop in the years before the final menstrual period and may persist for a few or for many years in postmenopause.

Vasomotor Symptoms (VMS)
Vasomotor symptoms (VMS), hot flushes and night sweats, are the hallmarks of menopause, occurring in approximately 75 percent of postmenopausal women with 25% of these being severely affected. The percentage of women reporting hot flushes varies across countries and ethnic backgrounds. Symptoms may resolve in 2-5 years but the median duration is 7 years and sometimes longer (Avis, 2015).

Hot flushes often begin as the sudden sensation of heat centred on the upper chest and face. In some instances, this will become generalised, lasting for several minutes, and can be associated with profuse perspiration, palpitations or anxiety which may be very distressing particularly when this happens repeatedly during the day and at night. At night, hot flushes and night sweats will often cause insomnia and fatigue. The differential diagnosis includes several entities distinguishable by clinical features such as thyroid disorder. Flushes can be related to drugs that affect vascular reactivity, such as some antihypertensives as well as commonly prescribed antidepressants when administered at high doses (SSRIs). The mechanism of VMS appears to involve the central nervous system, possibly due to narrowing of the thermoregulatory-neutral zone in women with hot flushes, associated with instability of the skin blood vessels.

The most effective treatment for VMS has been considered to be hormone replacement therapy, since symptoms occur at a time when oestrogen levels are dropping and ‘replacement’ leads to relief. HRT comprises synthetic hormones that may be identical to those produced from the ovaries during the reproductive years (oestradiol and progesterone) although other similar compounds (such as conjugated equine oestrogens, oestradiol valerate and several synthetic progestogens) are widely used. Tibolone belongs to the group of normethyltestosterone progestogen derivatives; it has metabolites that exhibit estrogenic, progestogenic and androgenic effects, and has been in clinical use since the early 1990s for treatment of menopausal symptoms.

However, many women do not take HRT. Some women elect to take no treatment as VSM may resolve naturally. Some simply do not wish to take hormones, while for others HRT is contraindicated, for example women who have (or are at high risk of) hormone-dependent cancer. Some women with hormone-dependent breast cancer experience severe flushing in association with a common long-term treatment (tamoxifen). We consider available
alternatives for these women in this chapter, including SSRI/SNRI antidepressants that impact on neurotransmitters in the brain; gabapentin, also used for control of epileptic seizures and neuropathic pain; herbal preparations, isoflavones and non-drug therapies such as cognitive behaviour therapy, hypnosis and exercise. Herbal preparations, isoflavones and bioidentical hormones are unregulated and in many instances not subject to any quality control or research studies of sufficient power or quality. They may not be safer than standard preparations, evidence on efficacy and side effects is incomplete.

Mood Changes

Depression and mood change is common at times of hormonal change such as during the menstrual cycle, after pregnancy and in the perimenopausal period. In some instances this may lead to clinical depression although this is not the focus of this part of the guideline. Women often complain of anxiety and feeling low around the time of the menopause, with mood swings and feeling frustrated. Women are often distressed at these changes and feel that they are out of character. Some women will feel much better with HRT particularly if the mood change is associated with fatigue due to VSM.

Musculoskeletal symptoms

Joint and muscle aches and pains are often reported by women in menopause. Specific treatment is not usually offered, but these symptoms could be associated with lack of ovarian hormone production and respond to HRT.

Sexual Disorders

Menopausal women may experience problems with sexual intercourse. This can be a complex issue that has both physical and psychological elements. The vaginal dryness resulting from urogenital atrophy can lead to pain with intercourse which can impact on libido. Loss of libido may also be a result of declining levels of oestrogen and testosterone as the ovaries fail; the lack of testosterone can be more marked in women who have their ovaries removed by surgery. Vaginal dryness tends to increase in severity with time since menopause. Topical treatment may be offered, both hormonal and non-hormonal.

The impact of severe menopausal symptoms on quality of life may be substantial and some women for whom HRT is contraindicated may choose to accept a degree of risk that might be considered by some to outweigh the benefits of MHT. A fully informed patient should be empowered and supported to make a decision that best balances benefits to that individual when weighed against potential risks.

7.3 Clinical introduction

This review question aims to assess the relative clinical effectiveness of the most common treatments used to relieve short term menopause-related symptoms for women. As this question was set out to assess the comparative effectiveness of all the main interventions, RCTs were selected as the best study design to answer this review question.

The main categories of interventions included in this review question were hormonal pharmaceutical treatments, non-hormonal pharmaceutical treatments, non-pharmaceutical treatments, and psychological therapies. The main short term menopausal symptoms that were the focus of this question were the following:

- frequency of vasomotor symptoms
- anxiety and low mood (excluding clinical depression) as aspects of psychological wellbeing. Depression in the context of this review question referred to low mood, as no clinical diagnosis was made and this term (low mood) will be used across the review.
- frequency of sexual intercourse as a measure of sexual function
• joint and muscle aches and pains as indicators of musculoskeletal symptoms

In order to capture the spectrum of adverse events that may be associated with different treatments used for the relief of menopausal related symptoms, vaginal bleeding and discontinuation of treatment due to side effects were selected as the most representative measures of women’s experience of adverse events in the short term. Long term adverse effects of HRT are covered in other sections (10-10.8).

The presentation of evidence synthesis is divided in 2 parts, based on the type of analysis which was used to produce these syntheses:

• a network meta-analysis (NMA) was conducted for the outcomes of vasomotor symptoms, vaginal bleeding and discontinuation. These outcomes were prioritised because they are highly prevalent among women who are seeking treatment for menopausal symptoms and due to their importance on continuity of health care and further impact on women’s experience of long term outcomes. 51 trials were included in the NMA for the outcomes of frequency of vasomotor symptoms, discontinuation and vaginal bleeding. Different number of trials contributed to each of NMA’s networks (ranging from 4 to 32 trials for each network).

• Pair-wise meta-analyses were conducted for the outcomes of low mood, anxiety, frequency of sexual activities and frequency of joint and muscle aches and pains. 69 trials were included in the pair-wise comparisons presenting these outcomes.

The NMA allows the synthesis of data from direct and indirect comparisons without breaking the randomisation of trials, in order to produce measures of class treatment effect and ranking of different interventions for the outcomes of interest. The NMA protocol was designed (please see full details in Appendix D) with the aim to provide a methodologically and clinically appropriate basis to address this review question. In summary, stratified analysis was pre-selected based on the 3 main groups of women in menopause; women with and without uterus and women with a history of breast cancer. For each of these strata, a list of the most appropriate interventions was organised; for example for women with a uterus the combination of oestrogen plus progestogen was selected as the most appropriate hormonal treatment because progestogen is needed in women with a uterus to prevent the proliferation of the endometrium which could cause endometrial cancer if not controlled. Only non-hormonal treatments were included for the group of women with a history of breast cancer due to the potential risk of cancer recurrence. A class effect model was selected for the NMA with the underlying assumption that the effectiveness of different treatments under the same class would be comparable. This decision was made in order to maximise the availability of data and borrow strength from different trials. Non hormonal treatments were common across the 3 strata. In addition, due to high variation in the way data was collected and presented in different trials in this area, we set up a clear and consistent approach of data collection. For example we decided a priori to examine the role of different treatments used for relief of vasomotor symptoms and not on their severity. Assumptions were also made for the minimum duration of trials for inclusion in the NMA and at the minimum acceptable criteria for mixed population studies. These assumptions are commonly made when a complex meta-analysis is designed and not only in the case of the NMA.

A number of studies (please see full details in Appendix G) were excluded from further analysis due to not meeting the minimum acceptable criteria when studies used mixed populations (therefore the interpretation of results would be confounded by the effect of differences in women’s baseline characteristics) and lack of information on the variation of estimate effects (no measures of SE, or SD were presented). The majority of reasons of exclusion of studies would also apply in a conventional pair-wise meta-analysis in order to produce reliable estimates of effects of different interventions. For the very minority of studies excluded for purely statistical reasons, their results were discussed with the GDG in relation to the interpretation of NMA results and whether the information of excluded studies would
change the direction of their decision making. This information was used as supplementary evidence to facilitate the GDG discussion.

For full details see review protocol in Appendix D.

### 7.4 Description of included studies

Description of included studies for the pair wise comparisons are given in Table 11.

#### Table 11: Evidence summary table for the pair wise comparisons for the outcomes of frequency of sexual function, anxiety and depression (not clinical)

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size per group</th>
<th>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</th>
<th>Outcomes</th>
<th>Preparation of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hormonal pharmaceutical treatments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davis 2008</td>
<td>Placebo (N = 277) Testosterone 150 ug/Day (N = 267) Testosterone 300 ug/day (N = 267) women were receiving concomitant oestrogen therapy</td>
<td>Surgical menopausal women: 20 - 70 years and postmenopausal for at least 12 months</td>
<td>Frequency of sexual activities</td>
<td>Testosterone 150 ug/Day, Testosterone 300 ug/day</td>
</tr>
<tr>
<td>Simon 2005</td>
<td>Placebo n=279 Testosterone n=283 (women were receiving concomitant oestrogen therapy)</td>
<td>20-70 year of age not at risk of breast or cervical cancer, have undergone bilateral salpingo-oophorectomy and hysterectomy at least 6 months before screening, and have no physical impediment to sexual function.</td>
<td>Frequency of sexual activities</td>
<td>Testosterone (300 mcg/d) or placebo patches applied twice weekly for 24 weeks</td>
</tr>
<tr>
<td><strong>Comparison of tibolone versus combined oestrogen/progesterone</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nijland 2008</td>
<td>Tibolone N=199 Transdermal E2/NETA N=201</td>
<td>Aged between 48 - 68 years - Undergone natural menopause, had intact uterus - Reported that prior to menopause, their sex life was satisfying but since menopause they experienced decline in satisfaction with sexual activity that was associated with personal distress as measured by Female Sexual Distress Scale (FSDS ≥ 15).</td>
<td>Frequency of sexual activities</td>
<td>E2 (50 ug)/NETA (140 ug) in the form of a twice weekly patch plus a daily placebo tablet - Tribolone 2.5 mg as a daily tablet with a twice weekly placebo patch.</td>
</tr>
<tr>
<td><strong>Hormonal pharmaceutical treatments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Nielsen 2006 | Intranasal 17B oestradiol: 150 ug/day: N = 114 300 ug/day: N = 103 Placebo: N = 118 | - 40 - 65 years old - Menopause defined as amenorrhea for more than 12 months or > 6 months with comitant serum level of oestradiol < 0.16 nmol/L plus FSH > 42 IU/L - All women who had undergone hysterectomy had menopause confirmed by determination of serum oestradiol and FSH at least 2 months prior to study entry. - Surgical menopause, if performed at least 6 weeks before study entry - Osteopenic (BMD T | Anxiety/low mood | Pulsed oestrogen therapy S21400 (intranasal 17B oestradiol): 150 ug/day and 300 ug/day or placebo - Women with intact uterus additionally received oral micronised progesterone 200 mg/day, 14 days out of 28
<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size per group</th>
<th>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</th>
<th>Outcomes</th>
<th>Preparation of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison of combined oestrogen with progesterone versus placebo</strong></td>
<td></td>
<td>Perimenopausal or postmenopausal intact uterus &gt;34 vasomotor symptoms (hot flashes and night sweats) per week Amenorrhea &gt;6 months and &lt;10 years FSH, &gt;40 mL/mL HT not contraindicated</td>
<td>Anxiety</td>
<td>Capsules were taken twice daily for 12 months -0.625 mg conjugated equine oestrogens plus 2.5 mg medroxyprogesterone acetate (CEE/MPA) (Placebo skin patch) 0.05 mg/day for 3 weeks</td>
</tr>
<tr>
<td>Geller 2009</td>
<td>Placebo arm: n = 22 randomised</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo arm: n = 21 included in analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oestrogens plus progestogen arm (CEE/MPA): n = 23 randomised and included in analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BC: n = 21 included in analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purdie 1995</td>
<td>HRT: 17</td>
<td>-Amenorrheic for at least 6 months - VSM symptoms - No HRT within past 6 months - Normotensive</td>
<td>Anxiety</td>
<td>- 0.625 mg CEE (regardless of hysterectomy status) plus 2.5 mg MPA or: - 0.625 mg CEE and 5 mg MPA if they were within 3 years from their last period</td>
</tr>
<tr>
<td></td>
<td>Placebo: 16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Veerus 2008</td>
<td>Blind HT arm: 415</td>
<td>-Aged 50 - 64 - Estonian speaking in 2 areas (Tallinn and Tartu) and in 2 counties surrounding these towns</td>
<td>Anxiety</td>
<td>0.625 mg CEE (regardless of hysterectomy status) plus 2.5 mg MPA or: - 0.625 mg CEE and 5 mg MPA if they were within 3 years from their last period</td>
</tr>
<tr>
<td></td>
<td>Placebo: N = 381</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-blind HT arm: N = 503</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-treatment arm: N = 524</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Veerus 2012</td>
<td>Non-HT arm (placebo and non-treatment arms): N = 673 HT arm (blind and non-blind HT arms): N = 686</td>
<td>-Aged 50 - 64 - Estonian speaking in 2 areas (Tallinn and Tartu)</td>
<td>Anxiety</td>
<td>0.625 mg CEE (regardless of hysterectomy status) plus 2.5 mg MPA or: - 0.625 mg CEE and 5 mg MPA if they were within 3 years from their last period</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comparison of oestrogen versus placebo</strong></td>
<td></td>
<td></td>
<td>Anxiety</td>
<td>0.625 mg / day CEE orally</td>
</tr>
<tr>
<td>Hachul 2008</td>
<td>CEE: 14</td>
<td>- Postmenopausal women - Aged 50 - 65 - Mean BMI less than 30 - No previous exposure to exogenous hormones</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Placebo: 19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schmidt 2000</td>
<td>34 female subjects, 16 received oestradiol first and 18 received placebo first</td>
<td>Self-report onset of depression associated with menstrual cycle irregularity of at least 6 months’ duration but with ≤1 of amenorrhea -diagnosis of major or minor depression determined by scores on the Centre for Epidemiologic Studies Depression Scale ≥10 during 3 of the 4 screening visits -plasma levels of follicle-stimulating hormone ≥20</td>
<td>Anxiety</td>
<td>Placebo skin patch for 3 weeks. 17β-oestradiol estraderm skin patch (0.05 mg/day) for 3 weeks</td>
</tr>
<tr>
<td>Study name</td>
<td>Sample size per group</td>
<td>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</td>
<td>Outcomes</td>
<td>Preparation of treatment</td>
</tr>
<tr>
<td>----------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Speroff 2003</td>
<td>Vaginal ring delivering 50 mcg per day E2 (n = 113) or 100 mcg per day E2 (n = 112), or a placebo vaginal ring (n = 108) for 13 weeks</td>
<td>At least 7 moderate to severe hot flushes per day or an average of at least 56 moderate to severe vasomotor symptoms per week for the 2 weeks before randomisation; if they did not have bilateral oophorectomy must have a FSH level of at least 40 IU and an E2 level of no more than 20 pg/mL</td>
<td>Anxiety</td>
<td>Vaginal ring delivering the equivalent of 50 mcg per day or 100 mcg per day of oestradiol or a placebo vaginal ring for 13 weeks</td>
</tr>
<tr>
<td>Thomson 1977</td>
<td>Oestrogen n=17 Placebo n=17</td>
<td>-Aged 45-55 -Amenorrhoea for at least 3 months -Symptoms of insomnia, depression, anxiety, and hot flushes</td>
<td>Anxiety</td>
<td>Piperazine oestrone sulphate in a dose of 1.5 mg twice daily; placebo</td>
</tr>
<tr>
<td>Comparison of oestrogen versus tibolone</td>
<td>Tibolone n=20 17 beta-oestradiol n=20</td>
<td>-Hysterectomy and bilateral oophorectomy -Perimenopausal period before the operation</td>
<td>Anxiety</td>
<td>Tibolone 2.5 mg/day or 17β- oestradiol 2 mg/day for 6 months</td>
</tr>
<tr>
<td>Wu 2001</td>
<td>Tibolone n=24;Continuous combined HRT (CEE plus MPA) n=24</td>
<td>12-36 months postmenopausal At least one climacteric symptom according to the Greene Climacteric Scale</td>
<td>Anxiety</td>
<td>Tibolone 2.5mg/day CEE 0.625 mg/day plus MPA 5mg/day Treatments were for 3 months</td>
</tr>
<tr>
<td>Comparison of testosterone versus placebo</td>
<td>Testosterone n=30 allocated, 3 discontinued Placebo n=30 allocated, 4 discontinued</td>
<td>Between 50 and 65 years of age and complaining of total loss or significant decrease of libido during the postmenopausal period</td>
<td>Anxiety</td>
<td>As a complement to their already on-going HRT (combined oestrogen and progesterone), 10 mg of a testosterone gel (Testogel, Besins–Iscovesco) or placebo was administered to the subjects.</td>
</tr>
<tr>
<td>Polisseni 2013</td>
<td>Tibolone (N = 42) E2 plus NETA (N = 44) Control (Ca plus Vit D3) (N = 44)</td>
<td>Between 45 - 60, postmenopausal with moderate - pronounced VSM symptoms &amp; Blatt-Kupperman Menopausal index (BKMI) equal to or greater than 20 Menopause characterised by the absence of menstruation for at least 12 months &amp; confirmed by increase of FSH</td>
<td>Anxiety</td>
<td>- 2.5 mg Tibolone -1mg oestradiol plus 0.5 mg norethindrone acetate - Control: 50 mg Calcium carbonate plus 200 UI vitamine D3</td>
</tr>
</tbody>
</table>
### Comparison of combined oestrogen with progesterone

**Zheng 2013**

- **Sample size per group**
  - Group A: Cimicifuga rhizome extract, n=32 (n=31 completed treatment)
  - Group B: Oestradiol valerate plus progesterone, n=32 (n=30 completed treatment)
  - Group C: Oestradiol valerate plus medroxyprogesterone acetate (MPA), n=32 (n=28 completed treatment)

- **Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)**
  - Women aged 40 to 60 years, early menopausal, going through climacteric symptoms
  - Early menopause was defined as going through amenorrhea above 6 months and within 5 years, serum E2 concentration <30pg/ml, and serum follicle stimulating hormone (FSH) concentration >40 IU/L

- **Outcomes**
  - Anxiety

- **Preparation of treatment**
  - Black cohosh
  - E2V plus Progesterone
  - E2V plus MPA

### Non-hormonal pharmaceutical treatments

#### Comparison of SNRI versus SSRI

**Soares 2010**

- **Sample size per group**
  - Acute Desvenlafaxine: 224
  - Escitalopram: 237
  - Continuation Phase Desvenlafaxine: 137
  - Escitalopram: 160

- **Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)**
  - Postmenopausal, between 40 - 70 years with primary diagnosis of MDD
  - Depressive symptoms for at least 30 days before screening visit and MADRS total score of 22 or higher

- **Outcome**
  - Anxiety

- **Preparation of treatment**
  - SNRI: desvenlafaxine 100-200 mg/day
  - SSRI: escitalopram 10-20 mg/day

### Comparison of gabapentin versus placebo

**Guttuso 2003**

- **Sample size per group**
  - Gabapentin n=30 assigned and analysed
  - Placebo n=29 assigned and analysed

- **Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)**
  - An average of 7 or more hot flashes per day accompanied by sweating
  - At least one daytime hot flash per day
  - Amenorrhea for more than 12 months or amenorrhea for 6–12 months with a serum follicle-stimulating hormone level greater than 40 mIU/mL and oestrogen less than 20 pg/mL or status post-bilateral oophorectomy for 2 months
  - An estimated creatinine clearance of 60 or more mL per minute
  - No oestrogen, progesterone, leuprolide, or tamoxifen therapy within the past 2 months
  - No change in dose of raloxifene, clonidine, or any antidepressant therapy within the past month and no plan to change the dose in the future
  - No calcium channel antagonist or gabapentin therapy within the past 2 weeks
  - No previous allergic reaction to gabapentin

- **Outcomes**
  - Anxiety

- **Preparation of treatment**
  - Gabapentin 900 mg per day or identically appearing placebo for 12 weeks

### Comparison of SSRI versus placebo

**Barton 2010**

- **Sample size per group**
  - 10 mg citalopram/placebo: n=44 / n=22

- **Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)**
  - Postmenopausal and reported to be bothered with at least 14 hot flashes per week for at least 1 week

- **Outcomes**
  - Anxiety

- **Preparation of treatment**
  - Citalopram at target doses of 10, 20, or 30 mg/d versus
### Menopause

**Managing short-term symptoms**

National Collaborating Centre for Women’s and Children’s Health

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size per group</th>
<th>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</th>
<th>Outcomes</th>
<th>Preparation of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>citalopram/placebo: n=44 / n=21 30 mg citalopram/placebo: n=44 / n=21</td>
<td>least the past month</td>
<td>placebo for 6 weeks.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Comparison of phytoestrogens versus placebo**

Evans 2011

Genistein n=42 assigned, n=40 intention-to-treat
Placebo n=42 assigned and intention-to-treat

Subjects had to have a minimum of 40 hot flushes per week, be between the ages of 40 and 65 and be in a physiological state of natural or surgical menopause

- Anxiety
- Placebo or a single 30 mg dose of synthetic genistein daily for 12 weeks

Geller 2009

Placebo arm: n = 22 randomised
Placebo arm: n = 21 included in analysis
Red clover arm (RC): n = 22 randomised and included in analysis

Perimenopausal or postmenopausal intact uterus
>34 vasomotor symptoms (hot flashes and night sweats) per week
Amenorrhoea >6 months and <10 years FSH, >40 mIU/mL
HT not contraindicated

- Anxiety
- Capsules were taken twice daily for 12 months
  - Red clover
  - Placebo

Tice 2003

Promensil n=84 assigned and analysed
Rimostil n=83 assigned and analysed
Placebo n=85 assigned and analysed

-45 to 60 years
-Experiencing at least 35 hot flashes per week
-Had a follicle-stimulating hormone (FSH) level of 30 mIU/mL
-Had either documented bilateral oophorectomy or at least 2 consecutive months of amenorrhoea prior to enrolment with at least 6 months of amenorrhoea in the year prior to entry

- Anxiety
- Promensil (82 mg of total isoflavones per day)
  - Rimostil (57 mg of total isoflavones per day)
  - Identical placebo contained less than 0.04 mg of total isoflavones per tablet
  - Participants were instructed to take 2 tablets once daily for 12 weeks

**Comparison of herbal preparations versus placebo**

Yang 2007

Pycnogenol (N = 80)
Placebo (N = 75)
No menopausal cycle for 3 - 11 months but normal cycles appeared again (perimenopausal)
- Hormone level FSH > 30 IU and oestrogen E2 < 20 pg/l

- Anxiety
- Pycnogenol 100 mg

Geller 2009

Placebo arm: n = 22 randomised
Placebo arm: n = 21 included in analysis
Black cohosh arm (BC): n = 22 randomised
BC: n = 21 included in analysis

Perimenopausal or postmenopausal intact uterus
>34 vasomotor symptoms (hot flashes and night sweats) per week
Amenorrhoea >6 months and <10 years FSH, >40 mIU/mL
HT not contraindicated

- Anxiety
- Capsules were taken twice daily for 12 months
  - Black cohosh
  - Placebo

Wiklund 1999

Placebo = 191
Ginseng = 193
Aged 45 - 65, without HRT for previous 2 months and with no bleeding during previous 6 months

- Anxiety
- Ginseng

Amsterdam 2009

Black cohosh extract n = 15
Women who were either postmenopausal for ≥ 12

- Anxiety
- Black Cohosh (2 x 32 mg)
## Study name

<table>
<thead>
<tr>
<th>Study Name</th>
<th>Sample size per group</th>
<th>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</th>
<th>Outcomes</th>
<th>Preparation of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo n = 13</td>
<td></td>
<td>months or peri menopausal (with amenorrhea lasting to 2 to 11 months in the preceding year) - Perimenopausal women were ≥ 40 years old and had no other demonstrable reason for their amenorrhea - Women with prior hysterectomy and uncertain menopausal status had a serum FSH level of ≥ 40 mlU/ml - Had a DSM IV Axis I diagnosis of Anxiety Disorder due to menopause that was ascertained via the Structured Diagnostic Interview for DSM IV</td>
<td></td>
<td>capsules daily) Placebo (2 x 100% rice powder daily) Both for 12 weeks</td>
</tr>
<tr>
<td>van Die 2009</td>
<td>St John's Wort and Chaste: N = 50. Placebo: N = 50</td>
<td>40 - 60 years, postmenopausal or perimenopausal, experiencing a minimum of 5 hot flushes/sweating episodes per day and scoring 20 plus on Greene Climacteric Scale. - Hysterectomised women over 53 and FSH &gt; 25 IU/L.</td>
<td>Anxiety</td>
<td>St John's Wort (H. perforatum) and Chaste tree/berry (V. agnus-castus).</td>
</tr>
</tbody>
</table>

### Comparison of herbal preparations versus combined oestrogen with progesterone

Zheng 2013  
N=96 participated in study  
Group A: Cimicifuga rhizome extract, n=32 (n=31 completed treatment)  
Group B: Oestradiol valerate plus progesterone, n=32 (n=30 completed treatment)  
Group C: Oestradiol valerate plus medroxyprogesterone acetate (MPA), n=32 (n=28 completed treatment)  
Women aged 40 to 60 years, early menopausal, going through climacteric symptoms  
Early menopause was defined as going through amenorrhea above 6 months and within 5 years, serum E2 concentration <30pg/ml, and serum follicle stimulating hormone (FSH) concentration >40 IU/L  
Anxiety  
Black cohosh E2Vplus Progestosterone E2VplusMPA

### Comparison of CBT versus usual care

Mann 2012  
Usual care n=49 randomised, 45 analysed  
CBT n=47 randomised, 43 analysed  
At least ten problematic HFNS per week (confirmed by a 2-week diary and a screening interview) for a duration of 2 months or more  
-Had completed medical treatment for breast cancer (surgery, radiotherapy, or chemotherapy), and had no evidence of other cancers or metastases  
-Women taking adjuvant endocrine treatment were eligible  
Anxiety  
Group CBT

### Comparison of herbal preparations versus tibolone

Qu 2009  
Gnl: N = 21  
Control (tibolone): N = 21  
Aged 40 - 60 with at least 6 consecutive months of  
Anxiety  
GengNianLe (Gnl, also...
<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size per group</th>
<th>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</th>
<th>Outcomes</th>
<th>Preparation of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>amenorrhea with serum oestradiol level &lt; 20 pg/mL and FSH &gt; 40 mIU/mL - minimum of 1 month of low mood, total HAMD score &gt; 20</td>
<td>Low mood</td>
<td>called perimenopausa l relieving formula, a defined formula of Chinese medicinal herbs, tibolone</td>
</tr>
<tr>
<td>Comparison of oestrogen versus placebo</td>
<td></td>
<td>50-90 years of age -postmenopausal at least 1 year with follicular stimulating hormone ≥ 40 mIU/mL for those within 5 years of menopause -Score ≥10 on the Centre for Epidemiologic Studies Depression Scale and 8-20 on the Hamilton Depression Scale -Meet DSM-IV criteria for major depression, dysthymia, or minor depression</td>
<td>8 weeks of treatment with oestradiol (.1 mg/day) or placebo</td>
<td></td>
</tr>
<tr>
<td>Morrison 2004</td>
<td>26 amenorrhea</td>
<td></td>
<td>Low mood</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sample oestradiol</td>
<td>(n = 31) or placebo (n = 26)</td>
<td>Low mood</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 mg/day;</td>
<td></td>
<td>0.625 mg / day CEE orally</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n = 31) or placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(n = 26)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hachul 2008</td>
<td>CEE: 14 Placebo: 19</td>
<td>- Postmenopausal women - Aged 50 - 65 - Mean BMI less than 30 - No previous exposure to exogenous hormones</td>
<td>Low mood</td>
<td>Placebo skin patch for 3 weeks, 17β-oestradiol estraderm skin patch (0.05 mg/day) for 3 weeks.</td>
</tr>
<tr>
<td>Schmidt 2000</td>
<td>34 female subjects, 16 received oestradiol first and 18 received placebo first</td>
<td>Self-report onset of depression associated with menstrual cycle irregularity of at least 6 months’ duration but with ≤1 of amenorrhea -diagnosis of major or minor depression determined by a structured diagnostic interview -scores on the Centre for Epidemiologic Studies Depression Scale ≥10 during 3 of the 4 screening visits -plasma levels of follicle-stimulating hormone ≥20 IU/L on 3 of 4 screening visits</td>
<td>Low mood</td>
<td></td>
</tr>
<tr>
<td>Speroff 2003</td>
<td>Vaginal ring delivering 50 mcg per day E2 (n = 113) or 100 mcg per day E2 (n = 112), or a placebo vaginal ring (n = 108) for 13 weeks</td>
<td>At least 7 moderate to severe hot flushes per day or an average of at least 56 moderate to severe vasomotor symptoms per week for the 2 weeks before randomisation -Women with uterus were required to have had amenorrhea for more than 12 months before randomisation; if she had amenorrhea for less than 12 but at least 6 months, she was also required to have a FSH level of at least 40 IU and an E2 level of no more than 20 pg/mL -Women with hysterectomy must had</td>
<td>Low mood</td>
<td>Vaginal ring delivering the equivalent of 50 mcg per day or 100 mcg per day of oestradiol or a placebo vaginal ring for 13 weeks.</td>
</tr>
<tr>
<td>Study name</td>
<td>Sample size per group</td>
<td>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</td>
<td>Outcomes</td>
<td>Preparation of treatment</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-----------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bilateral oophorectomy performed more than 6 weeks before randomisation; if they did not have bilateral oophorectomy, must have a FSH level of at least 40 IU and an E2 level of no more than 20 pg/mL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomson 1977</td>
<td>Oestrogen n=17</td>
<td>-Aged 45-55</td>
<td>Low mood</td>
<td>Piperazine oestrone sulphate in a dose of 1.5 mg twice daily; placebo</td>
</tr>
<tr>
<td></td>
<td>Placebo n=17</td>
<td>-Amenorrhoea for at least 3 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-Symptoms of insomnia, depression, anxiety, and hot flushes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Novaes Soares 2001</td>
<td>Oestradiol group n=25</td>
<td>(1) age between 40 and 55 years</td>
<td>Low mood</td>
<td>Transdermal patches of 17β-oestradiol (100 µg) or placebo for 12-week</td>
</tr>
<tr>
<td></td>
<td>Placebo group n=25</td>
<td>(2) history of menstrual cycle irregularity or amenorrhoea for less than 12 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) serum level of FSH greater than 25 IU/L (to document the gonadotropins’ attempt to stimulate the declining ovarian function and, therefore, to confirm the perimenopausal status as the cause of menstrual irregularities)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4) diagnosis of MDD, dysthymic disorder, or minor depressive disorder, according to DSM-IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison of tibolone, combined oestrogen/progesterone, and control group</td>
<td>Polisseni 2013</td>
<td></td>
<td>Low mood</td>
<td>2.5 mg Livial® (2.5 mg tibolone) oral tablets; 2/10 mg Femoston® (2 mg 17β-oestradiol sequentially combined with 10 mg dydrogesterone) oral tablets</td>
</tr>
<tr>
<td></td>
<td>Tibolone (N = 42)</td>
<td>Between 45 - 60 years; postmenopausal with moderate - pronounced VSM symptoms &amp; Blatt-Kuppermann Menopausal index (BKMI) equal to or greater than 20</td>
<td></td>
<td>2.5 mg Livial® (2.5 mg tibolone) oral tablets; 2/10 mg Femoston® (2 mg 17β-oestradiol sequentially combined with 10 mg dydrogesterone) oral tablets</td>
</tr>
<tr>
<td></td>
<td>E2 plus NETA (N = 44)</td>
<td>Menopause characterised by the absence of menstruation for at least 12 months &amp; confirmed by increase of FSH</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control (Ca plus Vit D3) (N = 44)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison of oestrogen versus tibolone</td>
<td>Somunkiran 2007</td>
<td>-Hysterectomy and bilateral oophorectomy - Perimenopausal period before the operation</td>
<td>Low mood</td>
<td>Tibolone 2.5 mg/day or 17β-oestradiol 2 mg/day for 6 months</td>
</tr>
<tr>
<td></td>
<td>Tibolone n=20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17 beta-oestradiol n=20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison of combined oestrogen with progesterone versus tibolone</td>
<td>Elfituri 2005</td>
<td>Healthy non- hysterectomised Libyan women naturally or surgically menopausal, with menopausal symptoms. In naturally menopausal women, it was at least 12 months since the last menstrual period (LMP) and at least 3 months after the bilateral oophorectomy in surgically menopausal women</td>
<td>Low mood</td>
<td>2.5 mg Livial® (2.5 mg tibolone) oral tablets; 2/10 mg Femoston® (2 mg 17β-oestradiol sequentially combined with 10 mg dydrogesterone) oral tablets</td>
</tr>
<tr>
<td></td>
<td>Tibolone n=50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17 beta - Oestradiol/dydrogesterone n=50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wu 2001</td>
<td>Tibolone</td>
<td>12-36 months</td>
<td>Low mood</td>
<td>Tibolone</td>
</tr>
</tbody>
</table>
### Study name | Sample size per group | Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer) | Outcomes | Preparation of treatment
--- | --- | --- | --- | ---
Continuous combined HRT (CEE plus MPA) n=24 | Postmenopausal At least one climacteric symptom according to the Greene Climacteric Scale | | 2.5mg/day CEE 0.625 mg/day plus MPA 5mg/day Treatments were for 3 months |
### Comparison of combined oestrogen with progesterone versus placebo

**Derman 1995**
- Sequential oestrogen / progesterone (Trisequens) = 40; Placebo = 42
- Women aged 40 - 60 years who complained of menopausal symptoms
- Low mood
- Sequential 17 beta-oestradiol and norethindrone acetate (Trisequens)

**Lin 2011**
- DRSP/E2 n=183 Placebo n=61
- -24 or more moderate to severe hot flushes over 7 consecutive days during the 3-week screening period
- -Intact uterus with endometrial thickness < 5 mm by transvaginal ultrasonography or normal endometrial biopsy if endometrial thickness was ≥ 5 mm
- -Last menstrual bleed ≥ 1 year before, or bilateral oophorectomy ≥ 6 weeks before, or last natural menstrual bleed ≥ 6 months (but <1 year) previously, with serum follicle stimulating hormone ≥ 40 mIU/ml
- -Negative urinary pregnancy test
- -Negative bilateral mammography result
- Low mood
- 2 mg drospirenone/1 mg oestradiol (DRSP/E2) versus placebo taken daily orally for 4 28-day cycles (16 weeks)

**Purdie 1995**
- HRT: 17 Placebo: 16
- -Amenorrhic for at least 6 months
- - VSM symptoms
- - No HRT within past 6 months
- - Normotensive
- Low mood
- HRT - 0.625mg conjugated equine oestrogen (orally), progesterone norgestrel 0.15 mg taken from days 17 - 28

**Rudolph 2004**
- 2 mg Oestradiol valerate (EV) plus 2 mg Dienogest n=65; Placebo n=64
- - Healthy postmenopausal women
- - 48 - 65 years
- - Mild to moderate depressive episode according to ICD10 and HAMD > 16
- Low mood
- 2 mg Oestradiol valerate (EV) plus 2 mg Dienogest (DNG) per day

**Veerus 2008**
- Blind HT arm: 415 Placebo: N = 381
- Non-blind HT arm: N = 503
- Non-treatment arm: N = 524
- -Aged 50 - 64
- - Estonian speaking in 2 areas (Tallinn and Tartu) and in 2 counties surrounding these towns
- Low mood
- - 0.625 mg CEE (regardless of hysterectomy status) plus 2.5 mg MPA or:
- - 0.625 mg CEE and 5 mg MPA if they were within 3 years from their last period

**Veerus 2012**
- Non-HT arm (placebo and non-treatment arms): N = 673
- HT arm (blind and non-blind HT arms): N = 524
- Aged 50 - 64
- - Estonian speaking in 2 areas (Tallinn and Tartu)
- Low mood
- 0.625 mg CEE (regardless of hysterectomy status) plus 2.5 mg MPA or:
<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size per group</th>
<th>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</th>
<th>Outcomes</th>
<th>Preparation of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison of testosterone versus placebo</td>
<td></td>
<td>Testosterone n=30 allocated, 3 discontinued Placebo n=30 allocated, 4 discontinued</td>
<td>Low mood</td>
<td>As a complement to their already on-going HRT (combined oestrogen and progesterone), 10 mg of a testosterone gel (Testogel, Besins–Iscovesco) or placebo was administered to the subjects.</td>
</tr>
<tr>
<td>Nathorst-Boos 2006</td>
<td></td>
<td>Between 50 and 65 years of age and complaining of total loss or significant decrease of libido during the postmenopausal period</td>
<td></td>
<td>0.625 mg CEE and 5 mg MPA if they were within 3 years from their last period</td>
</tr>
<tr>
<td>Comparison of combined oestrogen with progesterone</td>
<td></td>
<td>Healthy women with an intact uterus, had climacteric symptoms or ongoing HRT - Aged 52 or over</td>
<td>Low mood</td>
<td>-CE/MPA 0.625 mg/5 mg -E2/NETA 2 mg/1 mg</td>
</tr>
<tr>
<td>Odmark 2004</td>
<td>N = 246 - CE/MPA: N = 123 - E2/NETA: N = 123</td>
<td>Women aged 40 to 60 years, early menopausal, going through climacteric symptoms Early menopause was defined as going through amenorrhea above 6 months and within 5 years, serum E2 concentration &lt;30pg/ml, and serum follicle stimulating hormone (FSH) concentration &gt;40 IU/L</td>
<td></td>
<td>Black cohosh E2VplusProgestrone E2VplusMPA</td>
</tr>
<tr>
<td>Zheng 2013</td>
<td>N=96 participated in study Group A: Cimicifuga rhizome extract, n=32 (n=31 completed treatment) Group B: Oestradiol valerate plusprogesterone, n=32 (n=30 completed treatment) Group C: Oestradiol valerate plusmedroxyprogesterone acetate (MPA), n=32 (n=28 completed treatment)</td>
<td>Women aged 40 to 60 years, early menopausal, going through climacteric symptoms Early menopause was defined as going through amenorrhea above 6 months and within 5 years, serum E2 concentration &lt;30pg/ml, and serum follicle stimulating hormone (FSH) concentration &gt;40 IU/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-hormonal pharmaceutical treatments</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison of SSRI versus placebo</td>
<td></td>
<td>Postmenopausal and reported to be bothered with at least 14 hot flashes per week for at least the past month</td>
<td>Low mood</td>
<td>Citalopram at target doses of 10, 20, or 30 mg/d versus placebo for 6 weeks.</td>
</tr>
<tr>
<td>Barton 2010</td>
<td>10 mg citalopram/placebo: n=44 / n=22 20 mg citalopram/placebo: n=44 / n=21 30 mg citalopram/placebo: n=44 / n=21</td>
<td>Aged 18 and older with localised breast cancer and receiving adjuvant tamoxifen therapy -Had at least one hot flash per day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kimmick 2006</td>
<td>Sertraline n=33 assigned, 25 analysed Placebo n=29 assigned, 22 analysed</td>
<td>Aged 18 and older with localised breast cancer and receiving adjuvant tamoxifen therapy -Had at least one hot flash per day</td>
<td>Low mood</td>
<td>6 weeks of sertraline (50 mg each morning) versus placebo</td>
</tr>
<tr>
<td>Comparison of SNRI versus SSRI</td>
<td></td>
<td>Postmenopausal, between 40 - 70 years with primary diagnosis of MDD - Depressive symptoms for at least 30 days before screening visit and MADRS total score of 22 or higher</td>
<td>Low mood</td>
<td>SNRI: desvenlafaxine 100-200 mg/day SSRI: escitalopram 10-20 mg/day</td>
</tr>
<tr>
<td>Study name</td>
<td>Sample size per group</td>
<td>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</td>
<td>Outcomes</td>
<td>Preparation of treatment</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td><strong>Non-pharmaceutical treatments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Comparison of herbal preparations versus placebo</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiklund 1999</td>
<td>Placebo = 191</td>
<td>Aged 45 - 65, without HRT for previous 2 months and with no bleeding during previous 6 months</td>
<td>Low mood</td>
<td>Ginseng</td>
</tr>
<tr>
<td></td>
<td>Ginseng = 193</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amsterdam 2009</td>
<td>Black cohosh extract n = 15, Placebo n = 13</td>
<td>Women who were either postmenopausal for ≥ 12 months or peri menopausal (with amenorrhea lasting to 2 to 11 months in the preceding year)</td>
<td>Low mood</td>
<td>Black Cohosh (2 x 32 mg capsules daily) Placebo (2 x 100% rice powder daily) Both for 12 weeks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Perimenopausal women were ≥ 40 years old and had no other demonstrable reason for their amenorrhea</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Women with prior hysterectomy and uncertain menopausal status had a serum FSH level of ≥ 40 mIU/ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Had a DSM IV Axis I diagnosis of Anxiety Disorder due to menopause that was ascertained via the Structured Diagnostic Interview for DSM IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Die 2009</td>
<td>St John's Wort and Chaste: N = 50, Placebo: N = 50</td>
<td>40 - 60 years, postmenopausal or perimenopausal, experiencing a minimum of 5 hot flushes/sweating episodes per day and scoring 20 plus on Greene Climacteric Scale. - Hysterectomised women over 53 and FSH &gt; 25 IU/L.</td>
<td>Low mood</td>
<td>St John's Wort (H. perforatum) and Chaste tree/berry (V. agnus-castus).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uebelhack 2006</td>
<td>Treatment (Black Cohosh): 151, Placebo: 143</td>
<td>45 - 60 years, experiences climacteric complaints with pronounced psychological component for at least 3 months, left untreated for at least 2 months - HAMD total score 15 - 23 points</td>
<td>Low mood</td>
<td>Black Cohosh 1 mg triterpene glycosides and St John's Wort extract (0.25 mg total hypericine) - Placebo 2 tablets orally twice per day (week 1 - 8) and 1 tablet orally twice per day (weeks 9 - 16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comparison of herbal preparations versus combined oestrogen with progesterone</strong></td>
<td></td>
<td>Women aged 40 to 60 years, early menopausal, going through climacteric symptoms Early menopause was defined as going through amenorrhea above 6 months and within 5 years, serum E2 concentration &lt;30pg/ml, and serum follicle stimulating hormone (FSH) concentration &gt;40</td>
<td>Low mood</td>
<td>Black cohosh E2VplusProges terone E2VplusMPA</td>
</tr>
<tr>
<td>Zheng 2013</td>
<td>N=96 participated in study</td>
<td>Group A: Cimicifuga rhizome extract, n=32 (n=31 completed treatment) Group B: Oestradiol valerate plusprogesterone, n=32 (n=30 completed treatment) Group C: Oestradiol valerate</td>
<td>Low mood</td>
<td></td>
</tr>
</tbody>
</table>
### Menopause

**Managing short-term symptoms**

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size per group</th>
<th>Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer)</th>
<th>Outcomes</th>
<th>Preparation of treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>plusmedroxyprogesterone acetate (MPA), n=32 (n=28 completed treatment)</td>
<td></td>
<td>IU/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comparison of acupuncture versus sham acupuncture</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bao 2014</td>
<td>Acupuncture n=25, analysed n=24  Sham acupuncture n=26, analysed n=23</td>
<td>-Postmenopausal breast cancer who had been receiving AI therapy for greater than or equal to 1 month  -Reported AI-associated musculoskeletal symptoms  -Had not received acupuncture within the past 12 months</td>
<td>Low mood</td>
<td>Acupuncture  Sham acupuncture</td>
</tr>
<tr>
<td><strong>Comparison of phytoestrogen versus placebo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evans 2011</td>
<td>Genistein n=42 assigned, n=40 intention-to-treat  Placebo n=42 assigned and intention-to-treat</td>
<td>Subjects had to have a minimum of 40 hot flushes per week, be between the ages of 40 and 65 and be in a physiological state of natural or surgical menopause</td>
<td>Low mood</td>
<td>Placebo or a single 30 mg dose of synthetic genistein daily for 12 weeks</td>
</tr>
<tr>
<td>Tice 2003</td>
<td>Promensil n=84 assigned and analysed  Rimostil n=83 assigned and analysed  Placebo n=85 assigned and analysed</td>
<td>-45 to 60 years  -Experiencing at least 35 hot flushes per week  -Had a follicle-stimulating hormone (FSH) level of 30 mIU/mL  -Had either documented bilateral oophorectomy or at least 2 consecutive months of amenorrhea prior to enrolment with at least 6 months of amenorrhea in the year prior to entry</td>
<td>Low mood</td>
<td>Promensil (82 mg of total isoflavones per day)  -Rimostil (57 mg of total isoflavones per day)  -Identical placebo contained less than 0.04 mg of total isoflavones per tablet  -Participants were instructed to take 2 tablets once daily for 12 weeks</td>
</tr>
<tr>
<td>de Sousa-Munoz 2009</td>
<td>Isoflavones extract (EG=experimental group) n=42  Placebo made of starch (CG=control group) n=42</td>
<td>Age from 45 to 60 years  -One year or more of amenorrhea for non-hysterectomised women  -The presence of vasomotor and depression symptoms clinically detectable  -Follicle-stimulating hormone (FSH) plasma levels greater than or equal to 25 IU/L</td>
<td>Low mood</td>
<td>Daily dose of 120 mg isoflavones divided into 2 oral doses of 60 mg; Control group received 2 daily doses of placebo (starch)</td>
</tr>
<tr>
<td><strong>Psychological therapies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comparison of CBT versus usual care</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mann 2012</td>
<td>Usual care n=49 randomised, 45 analysed  CBT n=47 randomised, 43 analysed</td>
<td>At least ten problematic HFNS per week (confirmed by a 2-week diary and a screening interview) for a duration of 2 months or more  -Had completed medical treatment for breast cancer (surgery, radiotherapy, or</td>
<td>Low mood</td>
<td>Group CBT</td>
</tr>
</tbody>
</table>
### Study name | Sample size per group | Inclusion criteria (age, peri or postmenopause, uterus presence, breast cancer) | Outcomes | Preparation of treatment
--- | --- | --- | --- | ---
Non-pharmaceutical treatments versus HRT
Comparison of herbal preparations versus tibolone
Qu 2009 | GNL: N = 21 Control (tibolone): N = 26 | Aged 40 - 60 with at least 6 consecutive months of amenorrhea with serum oestradiol level < 20 pg/mL and FSH > 40 mIU/mL - minimum of 1 month of low mood, total HAMD score > 20 | Low mood | GengNianLe (GNL, also called perimenopausa l relieving formula), a defined formula of Chinese medicinal herbs, tibolone

Non-hormonal pharmaceutical treatments versus HRT
Comparison of combined oestrogen with progesterone versus SSRI
Soares 2006 | Oestrogen and progestogen therapy (EPT) n=16 Escitalopram (ESCIT) n=16 | Perimenopausal and postmenopausal women, aged 40 to 60 years, who presented with depressive disorders and menopause-related symptoms | Low mood | 8 week open trial with ESCIT (flexible dose, 10-20 mg/day; fixed dose, 10mg/day for the first 4 weeks) or oestrogen plus progestogen therapy (ethinyl oestradiol 5 mcg/day plus norethindrone acetate 1 mg/day)

Joint and muscular pain and ache
Hormonal pharmaceutical treatments
Comparison of oestrogen versus placebo
Brunner 2010 | 5310 received conjugated equine oestrogens, 5429 assigned to placebo. | Postmenopausal women, aged 50 to 79 years at initial screening, were eligible if they had a prior hysterectomy and met specific health criteria (not reported in the study). | Joint and muscular pain and ache | 0.625 mg/day conjugated equine oestrogens (CEE= Premarin) or a matching placebo.

Comparison of combined oestrogen with progesterone versus tibolone
Psychological therapies
Comparison of CBT versus usual care
Mann 2012 | Usual care n=49 randomised, 45 analysed CBT n=47 randomised, 43 analysed | At least ten problematic HFNS per week (confirmed by a 2-week diary and a screening interview) for a duration of 2 months or more - Had completed medical treatment for breast cancer (surgery, radiotherapy, or chemotherapy), and had no evidence of other cancers or metastases - Women taking adjuvant endocrine treatment were eligible | Joint and muscular pain and ache | Group CBT
Table 12: Evidence summary table for studies included in the NMAs for the outcomes of vasomotor symptoms, discontinuation and vaginal bleeding, in women with uterus, without uterus, and with breast cancer/history of breast cancer

<table>
<thead>
<tr>
<th>Study name</th>
<th>Sample size per group</th>
<th>Description of treatment</th>
<th>Outcomes</th>
<th>Populations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-Akoum 2009</td>
<td>Placebo (N=25); St John's Wort (N=22)</td>
<td>Placebo (TID); Ethanolic St John's Wort extract, 900mg (300mg TID)</td>
<td>VMS</td>
<td>Women with a history of breast cancer</td>
</tr>
<tr>
<td>Al-Azzawi 1999</td>
<td>Oestradiol oral plus progesterone oral High (N=116); Tibolone Oral High (N=191)</td>
<td>2mg micronised oestradiol valerate and 0.7 mg norethisterone; 2.5mg/day tibolone</td>
<td>Bleeding</td>
<td>Uterus</td>
</tr>
<tr>
<td>Albertazzi 1998</td>
<td>Placebo (N=53); Isoflavones/Genistein/soy (N=51)</td>
<td>60g of placebo (casein) daily: 40g of proteins but no isoflavones: powder form in sachets of 30g each; 60g of isolated soy protein daily: contains 40g of proteins and 76mg of isoflavones (aglycone units) - powder form in sachets of 50g each</td>
<td>Discontinuation,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Baber 1999</td>
<td>Placebo (N=26); Isoflavones/Genistein/soy (N=25)</td>
<td>Placebo; 40mg/day phytoestrogen</td>
<td>VMS,</td>
<td>Women with a uterus, without uterus</td>
</tr>
<tr>
<td>Burke 2003</td>
<td>Placebo (N=70); Isoflavones/Genistein/soy (N=76); Isoflavones/Genistein/soy (N=65)</td>
<td>25 g of soy protein, alcohol washed to remove isoflavones (≤ 4 mg/day) (placebo); 25 g of soy protein with a medium dose of isoflavones (42 mg/day); 25 g of soy protein with a higher dose of isoflavones (58 mg/day)</td>
<td>VMS,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>D'Anna 2009</td>
<td>Placebo (N=191); Isoflavones/Genistein/soy (N=198)</td>
<td>Placebo; 54mg/day genistein</td>
<td>VMS,</td>
<td>Women with a uterus, without uterus</td>
</tr>
<tr>
<td>Endrikat 2007</td>
<td>Placebo (N=162); Oestradiol valerate plus oral progesterone Ave (N=162)</td>
<td>Placebo; 2mg dienogest/1mg oestradiol valerate</td>
<td>Discontinuation,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td>Evans 2010</td>
<td>Placebo (N=42); Isoflavones/Genistein/soy (N=42)</td>
<td>Placebo; 30mg/d genistein</td>
<td>Discontinuation,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td>Faure 2002</td>
<td>Placebo (N=36); Isoflavones/Genistein/soy (N=39)</td>
<td>2x2 capsules of placebo (cellulose microcrystalline/sodium magnesium stearic) per day; 2x2 capsules of soy isoflavone extract per day</td>
<td>VMS,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Ferrari 2009</td>
<td>Placebo (N=95); Isoflavones/Genistein/soy (N=85)</td>
<td>Placebo; 80mg/day phytoestrogen (corresponding to 60mg of genistein)</td>
<td>VMS, Discontinuation,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Freedman 2010</td>
<td>Placebo (N=12); 5-HTP (N=12)</td>
<td>Placebo; 150 mg of 5-hydroxytryptophan given daily</td>
<td>VMS,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Freedman 2011</td>
<td>Placebo (N=14); Citalopram (N=12)</td>
<td>Placebo; 10-20mg/day Escitalopram</td>
<td>VMS,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Freeman 2011</td>
<td>Placebo (N=101); Citalopram (N=104)</td>
<td>Placebo; 10 to 20 mg of escitalopram daily</td>
<td>VMS, Discontinuation,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Garcia 2010</td>
<td>Placebo (N=39); Multibotanicals (N=120)</td>
<td>Placebo; Mung legume extract combined with Eucommia ulmoides</td>
<td>VMS, Discontinuation,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Gordon 2006</td>
<td>Placebo (N=41); Sertraline (N=46)</td>
<td>Placebo; 50mg/day Sertraline</td>
<td>VMS,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Grady 2007</td>
<td>Placebo (N=49); Sertraline (N=50)</td>
<td>Placebo; 50mg/day Sertraline</td>
<td>VMS,</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Guttuso 2003</td>
<td>Placebo (N=29); Gabapentin (N=54)</td>
<td>Identically appearing placebo capsules; 900mg capsules of gabapentin/day</td>
<td>Discontinuation, Bleeding</td>
<td>Women with a uterus, women without uterus</td>
</tr>
<tr>
<td>Study name</td>
<td>Sample size per group</td>
<td>Description of treatment</td>
<td>Outcomes</td>
<td>Populations</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>---------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>Hachul 2011</td>
<td>Placebo (N=19);</td>
<td>Placebo; 80mg/day</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>Isoflavones/Genistein/</td>
<td>isoflavone</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>soy (N=19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hammar 2007</td>
<td>Tibolone High(N=285);</td>
<td>2.5 mg tibolone; 1 mg</td>
<td>Bleeding</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td></td>
<td>Oestradiol oral plus</td>
<td>17b oestradiol plus 0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>progestogen oral Ave</td>
<td>mg norethisterone acetate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=284)</td>
<td>daily for 48 weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joffe 2014</td>
<td>Placebo (N=146);</td>
<td>Placebo; Oestradiol oral</td>
<td>VMS, Discontinuation,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>Oestradiol oral plus</td>
<td>progestogen oral Low</td>
<td>Bleeding</td>
<td></td>
</tr>
<tr>
<td></td>
<td>progestogen oral</td>
<td>(0.5mg per day O plus</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low (N=96); Venlafaxine</td>
<td>10mg/day medroxyprogesterone if</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=97)</td>
<td>women had uterus)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Venlafaxine (37.5mg/day</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>for 1 week then 75mg/day</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>for 7 weeks)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kimmick 2006</td>
<td>Placebo (N=29);</td>
<td>Placebo; 50mg/day</td>
<td>VMS, Discontinuation,</td>
<td>Wome with a history of breast cancer</td>
</tr>
<tr>
<td></td>
<td>Sertraline (N=33)</td>
<td>sertraline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knight 1999</td>
<td>Placebo (N=12);</td>
<td>Placebo; 1 tablet (40 mg)</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>Isoflavones/Genistein/</td>
<td>of Promensil daily; 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>soy (N=12);</td>
<td>tablets (160 mg) of</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Isoflavones/Genistein/</td>
<td>Promensil daily</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>soy (N=12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knight 2001</td>
<td>Placebo (N=12);</td>
<td>Isosvalone-free, isocaloric</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>Isoflavones/Genistein/</td>
<td>casein-based beverage;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>soy (N=12)</td>
<td>Dietary beverage in the</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>form of soy powder</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>containing isoflavones, daily</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>dose of 4 scoops or 60g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landgren 2005</td>
<td>Placebo (N=58);</td>
<td>Placebo; Daily oral 1.25mg</td>
<td>VMS, Discontinuation,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td></td>
<td>Tibolone Low (N=73);</td>
<td>tibolone; Daily oral 2.5mg</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tibolone Ave (N=68);</td>
<td>tibolone; Daily oral 5.0mg</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tibolone High (N=57)</td>
<td>tibolone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lin 2011</td>
<td>Placebo (N=62);</td>
<td>Oral placebo once daily;</td>
<td>VMS,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td></td>
<td>Oestradiol oral plus</td>
<td>Oral 2mg drospirenone/1mg</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>progestogen oral Ave</td>
<td>oestradiol (DRSP/E2) once</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=187)</td>
<td>daily</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lipovac 2011</td>
<td>Placebo (N=60);</td>
<td>Placebo; 40mg red clover</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>Red clover (N=53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mirabi 2013</td>
<td>Placebo (N=38);</td>
<td>Placebo; Valerian root</td>
<td>Discontinuation,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>Valerian root (N=38)</td>
<td>(225mg, 3 times per day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nedeljkovic 2013</td>
<td>Placebo (N=10);</td>
<td>Placebo; Sham acupuncture</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>Sham acupuncture</td>
<td>(N=10); Acupuncture</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=17); Acupuncture</td>
<td>(N=12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nir 2007</td>
<td>Sham acupuncture</td>
<td>Placebo acupuncture, 9</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td></td>
<td>(N=17); Acupuncture</td>
<td>sessions twice weekly</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=12)</td>
<td>during the first 2 weeks,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>once weekly for the</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>remaining 5 weeks ;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Active acupuncture, 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notelovitz 2000</td>
<td>Placebo (N=53);</td>
<td>Transdermal placebo patch;</td>
<td>VMS,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td></td>
<td>Oestradiol transdermal</td>
<td>Transdermal patch 50mcg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>plus progestogen</td>
<td>oestradiol plus combination</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>transdermal Low</td>
<td>patch 50mcg/d oestradiol</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=55); Oestradiol</td>
<td>plus combination patch</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>transdermal plus</td>
<td>50mcg/d oestradiol plus</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>progestogen transdermal High</td>
<td>140 mcg/d of norethindrone acetate;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>transdermal Ave</td>
<td>Transdermal patch 50mcg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N=59); Oestradiol</td>
<td>oestradiol plus combination</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>transdermal plus</td>
<td>patch 50mcg/d oestradiol</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>progestogen transdermal High</td>
<td>plus combination patch 50mcg/d oestradiol plus 250 mcg/d of norethindrone acetate;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transdermal patch 50mcg/d</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>oestradiol plus combination</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study name</td>
<td>Sample size per group</td>
<td>Description of treatment</td>
<td>Outcomes</td>
<td>Populations</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Palacios 2004</td>
<td>Placebo (N=159); Raloxifene (N=161); Raloxifene (N=167)</td>
<td>Placebo; 60mg/day raloxifene (RLX); 60mg/day raloxifene every other day for 1st 2 months, followed by 60mg/d for remainder of study (SDE)</td>
<td>VMS, Discontinuation</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Panay 2009</td>
<td>Placebo (N=201); Oestradiol oral plus progesterone oral Low (N=194); Oestradiol oral plus progesterone oral Low (N=182)</td>
<td>Placebo; 0.5mg NETA plus 0.1mg oestradiol; 0.5mg NETA plus 0.25mg oestradiol</td>
<td>Discontinuation,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td>Pandya 2005</td>
<td>Placebo (N=137); Gaberpentin (N=144); Gaberpentin (N=139)</td>
<td>Placebo; 300mg/day gabapentin; 900mg/day gabapentin</td>
<td>VMS, Discontinuation,</td>
<td>Women with a history of breast cancer</td>
</tr>
<tr>
<td>Penotti 2003</td>
<td>Placebo (N=34); Isoflavones/Genistein/soy (N=28)</td>
<td>Two 0.5g of talc and 0.5g of microcrystalline cellulose placebo tablets per day (placebo); Two 72 mg of soy-derived isoflavones tablets per day</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Pinkerton 2009</td>
<td>Placebo (N=66); Bazedoxifene plus oestradiol (N=133); Bazedoxifene plus oestradiol (N=133)</td>
<td>Placebo; Bazedoxifene 20mg with conjugated oestrogen 0.45mg once daily; Bazedoxifene 20mg with conjugated oestrogen 0.625mg once daily</td>
<td>Discontinuation,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td>Pinkerton 2012</td>
<td>Placebo (N=190); Desvenlafaxine (N=200)</td>
<td>Placebo; Desvenlafaxine 100mg/d</td>
<td>Discontinuation,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Pinkerton 2013</td>
<td>Placebo (N=294); Gaberpentin (N=299)</td>
<td>Placebo; Gabapentin (600mg am/1200 mg pm)</td>
<td>Discontinuation,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Rotem 2007</td>
<td>Placebo (N=25); Black cohosh (N=25)</td>
<td>Placebo; Phyto-Female Complex (standardised extracts of black cohosh, dong quai, milk thistle, red clover, American ginseng, chastetree berry) daily</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Schurmann 2004</td>
<td>Placebo (N=61); Oestradiol oral plus progesterone oral Ave (N=57); Oestradiol oral plus progesterone oral Ave (N=55); Oestradiol oral plus progesterone oral Ave (N=52)</td>
<td>Placebo; 1mg oestradiol and 1mg drospirenone; oral tablet once daily; 1mg oestradiol and 2mg drospirenone; oral tablet once daily; 1mg oestradiol and 3mg drospirenone; oral tablet once daily</td>
<td>Discontinuation,</td>
<td>Women with a uterus</td>
</tr>
<tr>
<td>Shahnazi 2013</td>
<td>Placebo (N=42); Black cohosh (N=42)</td>
<td>Placebo; Black cohosh</td>
<td>VMS,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Speroff 1996</td>
<td>Placebo (N=52); Oestradiol alone transdermal Low (N=54); Oestradiol alone transdermal Low (N=53)</td>
<td>One placebo transdermal system applied weekly; Two placebo transdermal system applied weekly; One 7-day transdermal system which delivered 0.02mg of 17beta-oestradiol/day applied every week</td>
<td>Discontinuation,</td>
<td>Women without a uterus</td>
</tr>
<tr>
<td>Stearns 2013</td>
<td>Placebo (N=56); Paroxetine (N=58); Paroxetine (N=51)</td>
<td>Placebo; 12.5mg/d paroxetine; 25mg/d paroxetine</td>
<td>Discontinuation,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
<tr>
<td>Stevenson 2010</td>
<td>Placebo (N=127); Oestradiol oral plus progesterone oral Low (N=124); Oestradiol oral plus progesterone oral Ave (N=62)</td>
<td>Placebo; 0.5mg/2.5mg CEE daily; 1mg/5mg CEE daily</td>
<td>VMS, Discontinuation, Bleeding</td>
<td>Women without a uterus</td>
</tr>
<tr>
<td>van de Weijer 2002</td>
<td>Placebo (N=16); Isoflavones/Genistein/soy (N=16)</td>
<td>Placebo; 80 mg isoflavones</td>
<td>VMS, Discontinuation,</td>
<td>Women with a uterus, women without a uterus</td>
</tr>
</tbody>
</table>
### 7.5 Clinical Evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.

Please refer to Appendix K for full details on presentation of NMA design, results, quality assessment and discussion.

### 7.6 Economic evidence

9 health economic studies were included for a review of treatment for the relief of individual menopause-related symptoms for women at menopause. These studies included various HRT alternatives, tibolone and no therapy. All the studies found active treatment to be cost-effective against no therapy.

A US study (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. 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 per QALY relative to no therapy.

A Canadian study compared continuous combined therapy 1mg norethindrone acetate and 5mcg ethinyloestradiol (NA/EE) versus 0.625mg conjugated equine oestrogen and 2.5mg medroxyprogesterone acetate (CEE/MPA) versus no therapy (Coyle 2003). The authors concluded that NA/EE was the most cost-effective intervention with an ICER of 20,300 CAD per QALY relative to CEE/MPA as a first line treatment.

The cost-effectiveness of HRT therapy versus placebo was reassessed for an average population of Swedish women with menopausal symptoms (Zethraeus 2005). 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.

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 results showed that compared to the high dose treatment, the low dose 0.3/1.5mg CE/MPA is the most cost-effective treatment, dominating the high dose alternative.

A Canadian study (Brown 2006) compared the cost-effectiveness of transdermal HRT against oral HRT and against placebo for women with post-menopausal symptoms. 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 Finnish trial based economic evaluation (Ylikangas 2007) compared continuous combined therapies for women with menopausal symptoms with a control group using data from the general Finnish population. The authors reported 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.

A Canadian study compared 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 authors concluded that tibolone is a cost-effective alternative to CEE/MA with an ICER of 9198 CAD per QALY.

A UK economic evaluation 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). The authors reported that treatment with HRT for menopausal symptoms was cost-effective in both groups of women with ICERs of 580 GBP per QALY and 205 GBP per QALY respectively. 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 women. The authors reported that HRT was cost-effective in women with menopausal symptoms, with an ICER of 2803 USD per QALY in women with an intact uterus and an ICER of 295 USD per QALY in hysterectomised women. A fuller description of this review of this evidence is provided in Appendix L.

However, none of the included published health economic studies considered the range of treatment alternatives being addressed by this guideline and nor did they use the techniques of network meta-analysis to synthesise direct and indirect evidence of clinical efficacy. Therefore, a de Novo model was developed for this guideline using the results of the network meta-analysis also undertaken for this guideline.

The model was developed to compare 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. The model was developed for 3 populations, reflecting the populations used in the network met-analysis.

i. Women with a uterus
   - Interventions compared in this population
     - No treatment
     - Acupuncture
     - Chinese herbal medicine
     - Gabapentin
     - Isoflavones/Genisten/Soy
     - Multibotanicals
     - Oestradiol plus progestogen non-oral
     - Oestradiol plus progestogen oral
     - Black cohosh
ii. Women without a uterus

- Interventions compared in this population
  - No treatment
  - Acupuncture
  - Chinese herbal medicine
  - Gabapentin
  - Isoflavones/Genisten/Soy
  - Multibotanicals
  - Oestriadiol non-oral
  - Oestriadiol oral
  - Black cohosh
  - Valerian root
  - SSRIs/SNRIs

iii. Women with a history of breast cancer

- Interventions compared in this population
  - No treatment
  - Gabapentin
  - Isoflavones/Genisten/Soy
  - SSRIs/SNRIs
  - St John's Wort

The model included short term outcomes on hot flushes, bleeding (where appropriate) and discontinuation. It also included the impact of breast cancer for women and VTE as a result of up to 5 years use of HRT. The clinical reviews for the guideline were used to synthesise evidence on health effects with a network meta-analysis used to estimate the effects of treatment on vasomotor symptoms, discontinuation and bleeding. The pairwise meta-analysis undertaken for this guideline was used to estimate treatment effects for breast cancer and venous thromboembolism (VTE) outcomes. Health outcomes were translated into QALYs as part of a cost utility analysis with health state utility estimates for the various outcomes based on published estimates. These health state utilities are listed in Table 46 of Appendix L.

The analysis of costs was based on a NHS and personal social services (PSS) perspective with costs and effects discounted at a rate of 3.5% in line with the NICE reference case. The costs of treatment used in the analysis are shown in Table 35 in Appendix L. Other model costs are also described in Appendix L. A number of inputs were altered as part of a sensitivity analysis with all results presented being based on a probabilistic sensitivity analysis.

In women with a uterus the model suggested that transdermal HRT was cost-effective with this result becoming more pronounced with increasing symptom severity as measured by the mean number of hot flushes per day. The base case results are shown in Table 13. In general sensitivity analysis suggested this result would hold even with less favourable model inputs for transdermal HRT although the cost-effectiveness of transdermal HRT was more borderline if the reduction of hot flushes was assumed to have less impact on HRQoL within the model.

### Table 13: Base case results for women with a uterus based on 10,000 simulations*

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean cost</th>
<th>Mean QALY</th>
<th>Net mean benefit</th>
<th>Probability cost-effective</th>
<th>ICER</th>
</tr>
</thead>
</table>
### Table 14: Base case results for women without a uterus based on 10,000 simulations

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean cost</th>
<th>Mean QALY</th>
<th>Net mean benefit</th>
<th>Probability cost-effective</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td>No treatment</td>
<td>£0</td>
<td>0.0000</td>
<td>£0</td>
<td>2.3%</td>
<td>n/a</td>
</tr>
<tr>
<td>SSRIs/SNRIs</td>
<td>£34</td>
<td>0.0415</td>
<td>£797</td>
<td>18.6%</td>
<td>£813</td>
</tr>
<tr>
<td>Gabapentin</td>
<td>£52</td>
<td>0.0587</td>
<td>£1,122</td>
<td>14.9%</td>
<td>£1,042</td>
</tr>
<tr>
<td>Isoflavones/Genistein/Soy</td>
<td>£312</td>
<td>0.1089</td>
<td>£1,866</td>
<td>2.3%</td>
<td>Extended dominance</td>
</tr>
<tr>
<td>Oestradiol plus progestogen oral</td>
<td>£385</td>
<td>0.0784</td>
<td>£1,183</td>
<td>1.8%</td>
<td>Dominated</td>
</tr>
<tr>
<td>Valerian root</td>
<td>£437</td>
<td>0.0001</td>
<td>-£436</td>
<td>0.0%</td>
<td>Dominated</td>
</tr>
<tr>
<td>Black cohosh</td>
<td>£448</td>
<td>0.1646</td>
<td>£2,845</td>
<td>27.1%</td>
<td>£3,740</td>
</tr>
<tr>
<td>Multibotanicals</td>
<td>£483</td>
<td>0.0504</td>
<td>£524</td>
<td>5.0%</td>
<td>Dominated</td>
</tr>
<tr>
<td>Acupuncture</td>
<td>£545</td>
<td>0.1084</td>
<td>£1,624</td>
<td>6.4%</td>
<td>Dominated</td>
</tr>
<tr>
<td>Tibolone</td>
<td>£598</td>
<td>-0.0017</td>
<td>£1,019</td>
<td>0.0%</td>
<td>Dominated</td>
</tr>
<tr>
<td>Oestradiol plus progestogen non-oral</td>
<td>£888</td>
<td>0.1845</td>
<td>£2,801</td>
<td>19.3%</td>
<td>£22,165</td>
</tr>
<tr>
<td>Chinese herbal medicine</td>
<td>£2,009</td>
<td>-0.0018</td>
<td>-£2,044</td>
<td>0.0%</td>
<td>Dominated</td>
</tr>
</tbody>
</table>

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

In women without a uterus, transdermal HRT was also found to be the most cost-effective treatment although the impact of transdermal oestrogen patches on hot flushes was extrapolated from the effectiveness of oestrogen and progestogen patch in women with a uterus, which may have led to an over-estimation of treatment efficacy in this patient group. The base case results are shown in Table 14.

In women with breast cancer gabapentin, isoflavones, SSRIs, St John’s Wort and no therapy were compared. St John’s Wort was the most cost-effective option although at the lower end of symptom severity, gabapentin also had a high probability of being cost-effective. The base case results for this population are shown in Table xx.
### Table 15: Base case results for women with breast cancer based on 10,000 simulations*  

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean cost</th>
<th>Mean QALY</th>
<th>Net mean benefit</th>
<th>Probability cost-effective</th>
<th>ICER</th>
</tr>
</thead>
<tbody>
<tr>
<td>No treatment</td>
<td>£0</td>
<td>0.0000</td>
<td>£0</td>
<td>9.3%</td>
<td>n/a</td>
</tr>
<tr>
<td>Gabapentin</td>
<td>£28</td>
<td>0.0598</td>
<td>£1,168</td>
<td>52.9%</td>
<td>£474</td>
</tr>
<tr>
<td>SSRIs/SNRIs</td>
<td>£33</td>
<td>-0.1662</td>
<td>-£3,358</td>
<td>2.8%</td>
<td>Dominated</td>
</tr>
<tr>
<td>Isoflavones/Genistein/Soy</td>
<td>£263</td>
<td>-0.0337</td>
<td>-£938</td>
<td>2.3%</td>
<td>Dominated</td>
</tr>
<tr>
<td>St John’s Wort</td>
<td>£459</td>
<td>0.0919</td>
<td>£1,379</td>
<td>32.7%</td>
<td>£13,435</td>
</tr>
</tbody>
</table>

*Mean costs and QALYs calculated relative to no treatment

This model is described in more detail in Appendix L.

### 7.7 Evidence statements

#### 7.7.1 Evidence summary from the NMA

32 RCTs of 12 treatment classes (placebo, sham acupuncture, oestrogen plus progestogen non-oral, oestrogen plus progestogen oral, tibolone, raloxifene, SSRIs/SNRIs, isoflavones, Chinese herbal medicine, black cohosh, multibotanicals, acupuncture) were included for the NMA for vasomotor symptoms in women with a uterus. The quality of the evidence was low due to high heterogeneity although no inconsistency was identified in the network. One included RCT was at very high risk of bias and ten were high risk. The other 21 RCTs were low or moderate risk. The results demonstrated a highly beneficial effect of non-oral oestradiol plus progestogen for relieving the frequency of vasomotor symptoms. Oral oestradiol plus progestogen may also be beneficial, though there was a degree of uncertainty regarding its efficacy. Isoflavones showed some efficacy when compared to placebo, though non-oral oestradiol plus progestogen gave significantly greater improvement in vasomotor symptoms when compared to this treatment. Black cohosh showed efficacy compared to placebo. However, results for isoflavones and black cohosh, as well as for multibotanicals and Chinese herbal medicine, should be interpreted with caution as the variety of herbal preparations used in studies may differ significantly.

21 RCTs of ten treatment classes (placebo, oestrogen plus progestogen oral, bazedoxifene plus oestrogen, tibolone, SSRIs/SNRIs, gabapentin, isoflavones, Chinese herbal medicine, multibotanicals, valerian root) were included for the NMA for discontinuation of treatment in women with uterus. Low quality evidence due to high heterogeneity within the network demonstrated that women treated with non-oral oestradiol plus progestogen or with bazedoxifene plus oestradiol were less likely to discontinue treatment than if they were treated with placebo or tibolone. However, those treated with SSRIs/SNRIs were more likely to discontinue treatment compared to placebo, as would be expected due to the serious side-effects profile of these treatments. Inconsistency could not be assessed in this network as there were no closed-treatment loops. Only four RCTs were high risk of bias. The other 17 were low or moderate risk.

5 RCTs of five treatment classes (placebo, oestrogen plus progestogen oral, tibolone, SSRIs/SNRIs, gabapentin) were included for the NMA for vaginal bleeding in women with a uterus. Neither heterogeneity nor inconsistency could be assessed in the network, as a fixed effects model was used and there were no closed-treatment loops. One study was at high risk of bias, one was low risk, and the other three were moderate risk. The sparseness of data within the network meant that there was a high degree of uncertainty in estimates, and no conclusions could be draw regarding effects of treatments on vaginal bleeding (adverse event).
The network on frequency of vasomotor symptoms (32 RCTs of nine treatment classes (placebo, sham acupuncture, raloxifene, SSRIs/SNRIs, isoflavones, Chinese herbal medicine, black cohosh, multibotanicals, acupuncture)) for women without a uterus did not include the hormonal treatment of oestrogen alone, as the relevant trials were excluded on the basis of either mixed population or lack of information on variation of effect estimates. Therefore, the final model included only non-hormonal and non-pharmaceutical treatments that restricted the generalisation and applicability of its results given that treatment of oestrogen alone is the current most common treatment offered to menopausal women without a uterus. Therefore the GDG decided not to consider the results of this network for decision making, given the limitation of their generalisability in the clinical context.

15 RCTs of eight treatment classes (placebo, oestrogen alone non-oral, SSRIs/SNRIs, gabapentin, isoflavones, Chinese herbal medicine, multibotanicals, valerian root) were included in the NMA for discontinuation of treatment in women without a uterus. Neither heterogeneity nor inconsistency could be assessed in the network, as a fixed effects model was used and there were no closed-treatment loops. Only three RCTs were high risk of bias. The other 12 were either low or moderate risk. Patients treated with SSRIs/SNRIs were more likely to discontinue treatment than those treated with placebo. There was a high degree of uncertainty in other estimates within the network.

4 RCTs of five treatment classes (placebo, SSRIs/SNRIs, gabapentin, isoflavones, St John’s Wort) were included for the NMA for vasomotor symptoms in women with breast cancer/history of breast cancer. The evidence was of moderate quality due to moderate heterogeneity within the network. However, the sparseness of data within the network meant that there was a high degree of uncertainty in estimates, and no conclusions could be drawn regarding efficacy of treatments for vasomotor symptoms. Inconsistency would not be assessed as there were no closed-treatment loops. Of the four RCTs included, two were at moderate risk of bias and two were low risk.

3 RCTs of four treatment classes (placebo, SSRIs/SNRIs, gabapentin, isoflavones) were included for the NMA for discontinuation of treatment in women with breast cancer/history of breast cancer. Neither heterogeneity nor inconsistency could be assessed in the network, as a fixed effects model was used and there were no closed-treatment loops. Two of the RCTs were at moderate risk of bias and one was low risk. The sparseness of data within the network meant that there was a high degree of uncertainty in estimates, and no conclusions could be drawn regarding discontinuation of treatment in women with breast cancer/history of breast cancer.

### 7.7.2 Evidence summary from the pair-wise comparisons

#### Comparison of oestrogen versus no treatment/placebo

**Anxiety**

Evidence from 1 RCT (n=34) showed no significant difference in anxiety in menopausal women who received oestrogen compared with those who received placebo at 2-month follow-up. The evidence was of very low quality.

Evidence from 1 RCT (n= 221) showed a significantly greater reduction in anxiety in menopausal women who received oestradiol either in a dosage of 50 or 100 mcg/day compared with those who received placebo at 13-week follow-up. The evidence was of moderate quality.

Evidence from 1 RCT (n=33) showed no significant difference in prevalence of self-reported anxiety in menopausal women who received oestrogen compared with those who received placebo. The evidence was of very low quality.
Menopause
Managing short-term symptoms

Low mood

Very low quality evidence from two RCTs (n=68) showed no significant difference in low mood in menopausal women who received oestrogen compared with those who received placebo.

Evidence from 1 RCT (n=50) showed a significant reduction in low mood in menopausal women who received oestrogen compared with those who received placebo at 8 or 12-week follow-up. The evidence was of low to moderate quality. The same study showed no significant reduction at 4-week follow-up and the evidence was of low quality.

Evidence of moderate quality from 1 RCT (n=232) showed no significant difference in anxiety and low mood in menopausal women who received oestradiol 150 µg/day compared with those who received placebo at 2-year follow-up but a significant increase in anxiety and low mood with a higher oestradiol dosage of 300. The evidence was of low to moderate quality.

Evidence from one RCT (n=57) showed no significant difference in low mood in menopausal women who received oestrogen compared with those who received placebo at 8-week follow-up. The quality was of low quality.

Musculoskeletal symptoms

Evidence from 1 large RCT (n=6,594) showed no significant difference in the risk for musculoskeletal symptoms in menopausal women who received oestrogen without joint pain at enrolment compared with those who received placebo at 1-year follow-up. The evidence was of moderate quality.

Moderate quality evidence from the same large RCT (n=2,987) showed no significant difference in the risk for musculoskeletal symptoms in menopausal women who received oestrogen with joint pain at enrolment compared with those who received placebo at 1-year follow-up.

Comparison of oestrogen plus progestogen verses no treatment/placebo

Anxiety

Evidence from 3 RCTs (n = 1,480) showed no significant difference in anxiety scores in menopausal women who received oestrogen combined with progestogen compared with those who received placebo. The evidence was of low quality.

Evidence from 1 RCT (n = 44) showed no significant difference in anxiety in menopausal women who received CEE plus MPA compared with those who received placebo at 12 months. The evidence was of moderate quality.

Low mood

Evidence from 5 RCTs (n = 1,691) showed a significantly greater reduction in low mood in menopausal women who received oestrogen combined with progestogen compared with those who received placebo (no information on follow-up). The evidence was of very low quality.
Evidence from 1 RCT (n = 128) showed a significantly greater reduction in depression in menopausal women who received oestrogen combined with progestogen compared with those who received placebo at 24 week follow-up. The evidence was of moderate quality.

**Comparison of tibolone verses no treatment/placebo**

**Anxiety and low mood**

One RCT (n = 86) found no significant difference in final anxiety and low mood scores in menopausal women who received tibolone compared with those who did not receive the HRT (12 month follow-up). The quality of the evidence for this outcome was low to high.

**Comparison of Testosterone verses no treatment/placebo**

**Frequency of sexual intercourse**

One RCT (n = 562) found a significant increase in frequency of sexual activities at 24 week follow-up in menopausal women who received testosterone compared with those who did not receive testosterone. The quality of the evidence for this outcome was low.

Moderate quality evidence from 1 RCT (n = 519) found a significant increase in the frequency of sexual activity at 4-week follow-up in menopausal women who received testosterone compared with those who did not receive testosterone.

Both studies reporting results for the outcome of frequency of sexual intercourse included the majority of women with surgical menopause.

**Low mood**

One RCT (n = 53) found no significant difference in final low mood scores in menopausal women who received testosterone compared with those who did not receive HRT. The quality of the evidence for this outcome was low.

**7.7.2.1 Comparison of different interventions versus other treatment (not placebo)**

**Tibolone versus conjugated equine oestrogens plus medroxyprogesterone acetate (CEE plus MPA)**

**Anxiety and low mood**

Very low quality evidence from a RCT with 36 menopausal women found no significant difference in change scores for either anxiety or low mood at 3 months follow-up in those women who received tibolone compared with those who received CEE plus MPA.

**CEE/MPA versus E2/NETA (both hormonal treatments)**

**Low mood**

High quality evidence from a RCT with 246 menopausal women found a significant reduction in low mood at 1 month for those who received CEE/MPA compared with those who received E2/NETA.
Oestradiol/Progesterone versus Oestradiol/MPA (both hormonal treatments)

**Anxiety/low mood**

Low and very low quality evidence from 1 RCT study \((n = 58)\) found no significant difference in final anxiety or low mood scores at 3 months in menopausal women who received oestradiol/progesterone compared with those who received oestradiol/MPA.

SSRI (non-hormonal pharmaceutical treatment) versus Oestrogen/Progestogen (hormonal treatment)

**Low mood**

One RCT with 32 menopausal women found a significant reduction in low mood at 8 weeks in menopausal women who received SSRI compared with those who received oestrogen plus progestogen. The quality of this evidence was low.

SNRI versus SSRI (both non-hormonal treatments)

**Anxiety/Depression (non-clinical)**

Low and moderate quality evidence from 1 RCT with 234 participants found no significant reduction in either anxiety or low mood at 8 months in women who received SNRI compared with those who received SSRI.

Tibolone versus Oestrogen plus Progestogen (E2/NETA)

**Frequency of sexual function**

Low quality evidence from a RCT with 400 women in menopause found no significant difference in frequency of sexual activity at 4-week follow-up in women who received combined E2/NETA compared with those who received tibolone.

Tibolone versus oestradiol

**Anxiety**

One RCT study \((n = 40)\) found no significant difference in final anxiety scores at endpoint in menopausal women who received tibolone compared with those who received oestradiol. The quality of the evidence for this outcome was very low.

Herbal treatment versus Oestrogen plus Progestogen

**Anxiety/low mood**

Low quality evidence from 1 RCT study \((n = 61)\) found no significant difference in final anxiety or low mood at 3 months in menopausal women who received black cohosh compared with those who received oestradiol plus progestogen.

Herbal treatment versus Oestrogen plus MPA

**Anxiety/low mood**

Low quality evidence from 1 RCT \((n = 59)\) found no significant difference in final anxiety or low mood scores at 3 months in menopausal women who received black cohosh compared with those who received oestradiol/MPA.
7.7.2.2 Comparison of non-hormonal pharmacological treatments versus no treatment

In relation to the relative effectiveness of non-hormonal pharmacological treatments to reduce anxiety or low mood compared to placebo, the following conclusions were made:

- moderate quality evidence from 1 RCT with 248 women in menopause comparing different dosages of citalopram and placebo demonstrated that 20mg citalopram was significantly more effective than placebo on reducing anxiety at 6 weeks, but no such an effect was found for the other dosages of citalopram (10mg and 30mg). No significant difference was found between these treatment dosages (10, 20 and 30 mg) of citalopram and placebo on low mood.

- setraline and gabapentin were found to be not significantly better than placebo in reducing low mood and anxiety respectively for menopausal women (very low and moderate quality evidence from two RCTs of less than 60 women in each trial).

Herbal treatments compared with placebo

None of the herbal treatments (ginseng, black cohosh, black cohosh plus St. John’s Wort, St. John’s Wort plus Chaste, pycogneal) included in the evidence basis was found to be significantly better than placebo on reducing either anxiety or low mood for menopausal women. The quality of this evidence ranged from moderate to very low quality for six RCTs with over 100 menopausal women.

Phytosterrogen treatments compared with placebo

In relation to the relative effectiveness of phytosterrogen treatments to reduce anxiety or low mood compared to placebo, the following conclusions were made:

- Promensil (82 mg), rimostil (57 mg), genistein (30 mg) and 120 mg of soy isoflavones extract were found no better than placebo to improve these outcomes at 12 weeks follow-up (moderate to very low quality evidence from three RCTs with sample sizes over 100 menopausal women)

- genistein (30 mg) and red clover (120 mg) were found more effective in significantly reducing anxiety in menopausal women compared to placebo (moderate quality evidence from 2 RCTs of 84 and 43 women respectively).

Acupuncture compared with sham acupuncture (placebo)

One RCT study (n = 47) found no significant difference in the changes of low mood at 8 weeks follow-up in menopausal women who received acupuncture compared with those who received sham acupuncture. The quality of the evidence for this outcome was moderate.

7.7.2.3 Comparison of psychological treatments versus usual care

Anxiety/Low mood

Moderate quality evidence from 1 RCT with 88 menopausal women comparing cognitive behavioural therapy (CBT) and usual care demonstrated that CBT was significantly more effective than usual care to reduce anxiety and low mood at 26 weeks follow-up.

7.7.2.4 Economic evidence

Original health economic analysis conducted for the guideline suggests that transdermal oestradiol and progestogen was the most cost effective treatment in women with a uterus and that cost effectiveness increased with severity of vasomotor symptoms (ICER: £22,165 per QALY for a mean of 3 hot flushes per day). The analysis was assessed as applicable with minor limitations.
Menopause
Managing short-term symptoms

Original health economic analysis conducted for the guideline suggests that non-oral oestradiol was the most cost effective treatment in women without a uterus and that cost effectiveness increased with severity of vasomotor symptoms (ICER: £2,149 per QALY for a mean of 3 hot flushes per day). The analysis was assessed as partially applicable with serious limitations due to the extrapolation of effectiveness from a different intervention in a different population.

Original health economic analysis conducted for the guideline suggests that St John’s Wort was the most cost effective treatment in women with breast cancer (ICER: £13,435 per QALY for a mean of 3 hot flushes per day). The analysis was assessed as applicable with minor limitations.

One cost utility analysis found that 1mg of norethindrone acetate/5 µg of ethinyl estradiol (NA/EE) was cost-effective compared to 0.625 mg/day of conjugated estrogens plus 2.5 mg of medroxyprogesterone (CEE/MPA) (ICER: dominant) and no therapy for the management of vasomotor symptoms (ICER: 6,200 USD per QALY). This analysis was assessed as partially applicable with minor limitations.

One cost utility analysis found that low dose 0.3mg conjugated oestrogen and 1.5mg medroxyprogesterone acetate injection (0.3/ 1.5mg CE/MPA) was cost-effective compared to higher dose 0.625mg conjugated oestrogen and 5mg medroxyprogesterone acetate injection (0.625/ 5mg CE/MPA) (ICER: dominant) in women with an intact uterus and menopausal symptoms. This analysis was assessed as partially applicable with minor limitations.

One cost utility analysis found that HRT therapy was cost-effective compared to placebo (ICER: 12,807 SEK per QALY) for women with a uterus and with menopausal symptoms. This analysis was assessed as partially applicable with minor limitations.

One cost utility analysis found that HRT therapy was cost-effective compared to placebo (ICER: 8,266 SEK per QALY) for women who had received a hysterectomy and with menopausal symptoms. This analysis was assessed as partially applicable with minor limitations.

One cost utility analysis found that oral HRT was cost-effective when compared with transdermal (ICER: dominant) for women with post-menopausal symptoms. This analysis was assessed as partially applicable with major limitations.

One cost utility analysis found that continuous combined HRT was cost-effective when compared to a control group in general population (ICER: 4613 Euros per QALY) for up to 9 years with an for non-hysterectomised women predominantly in the age range of 55-64 years who are experiencing climacteric symptoms. This was assessed as partially applicable with major limitations.

One cost utility analysis found that synthetic hormone tibolone was cost effective when compared to conjugated equine oestrogens 0.625mg with medroxyprogesterone acetate 2.5mg (CEE/MA) (ICER: 9198 CAD per QALY) in women with menopausal symptoms. This was assessed as partially applicable with major limitation.

One cost utility analysis found that combined 1 mg estradiol and 0.5 mg norethisterone was cost-effective when compared to no therapy (ICER: of £580 per QALY) for the treatment of menopausal symptoms in women with an intact uterus. This was assessed as applicable with major limitations.
One cost utility analysis found that estradiol alone was cost-effective when compared to no therapy (ICER: of £205 per QALY) for the treatment of menopausal symptoms in hysterectomised women. This was assessed as partially applicable with major limitations.

### 7.8 Evidence to recommendations

#### 7.8.1 Relative value placed on the outcomes considered

The selection of the most important outcomes (short term symptoms) for this review question was based on the high prevalence of these symptoms in peri and post-menopausal women which impacts on their overall quality of life. In considering menopause-related symptoms, the most important outcomes were selected as frequency of vasomotor symptoms, low mood (non-clinical depression), anxiety and frequency of sexual intercourse. Vasomotor symptoms (hot flushing and/or night sweats) were selected as the most critical outcomes as these are the most common type of symptoms among menopausal women in the UK and the most frequent reason for seeking medical advice. Flushing can have significant impact on women’s quality of life in terms of lost sleep, ability to function in everyday activities and a negative impact on their social and professional life. Frequency of flushing was decided as a measure for this outcome rather than severity, since the latter was not as widely reported and diverse scales have been used that would make the synthesis of evidence problematic and less precise. Hot flushing and night sweats were considered as a single entity for studies reporting them separately, given that they reflect women’s overall experience of frequency of vasomotor symptoms and are considered manifestations of the same underlying clinical problem.

In relation to adverse events, owing to wide variation and individualised experience of isolated adverse events, vaginal bleeding and discontinuation of treatment were selected as the most representative indicators of experience of adverse events. More specifically, discontinuation was selected as a proxy of tolerability, which may also reflect partly an aspect of treatment efficacy. Vaginal bleeding was selected as it was measurable and, if persistent, leads to further clinical investigation with considerable costs involved. Unscheduled vaginal bleeding is a common side effect with combined hormone therapy, but if persistent beyond 3 months then investigation may be required depending on the degree of clinical concern.

The selection of outcomes for inclusion in the NMA was based on both their clinical importance and relevance to patients. Frequency of vasomotor symptoms, discontinuation and vaginal bleeding were prioritised for inclusion in the NMA due to availability of data that allowed the formulation of networks.

Evidence on frequency of sexual intercourse, low mood (non-clinical depression), anxiety and frequency of joints pains and muscle aches was presented only in pair-wise meta-analyses when data were available.

The Guideline Committee discussed that these short term menopausal symptoms can either cluster or be interdependent for women in menopause, although for the purposes of presentation of evidence synthesis these outcomes were presented separately.

#### 7.8.2 Consideration of clinical benefits and harms

Since symptoms adversely affect quality of life for women in menopause, available treatments were evaluated to determine the balance of efficacy and tolerability.

The Committee acknowledged that the choice of treatment for the relief of short term menopausal symptoms is influenced by a number of factors, such as women’s personal choices and preferences, individual risk profile including comorbidities, and the level of information women receive about HRT and its impact on longer term outcomes (such as
cardiovascular disease, breast cancer, and osteoporosis). In addition, given that menopause is a transitional phase and women’s experience of symptom often changes, the Committee wanted to draw the attention to the attentions of health care professionals that they need to accommodate to these changes when planning the treatment for these symptoms or personal circumstances.

Two of the most critical factors determining the choice and appropriateness of treatments for relief of menopausal symptoms are whether a woman has a uterus or not, and whether she has a history of breast cancer (or other hormone sensitive conditions or other contraindications to hormone therapies) since the side-effect profiles of hormonal treatments are different for these groups of women.

For these reasons, evidence from the NMA was synthesized separately for women with and without a uterus and for women with a history of breast cancer. More specifically, for women with a uterus, oestrogen plus progestogen (transdermal) was found to be the most effective treatment to relieve vasomotor symptoms, with a significantly lower discontinuation rate compared to all the other available treatments (hormonal, non-hormonal and non-pharmacological). There was also some evidence that oestrogen plus progestogen (oral) may be more effective to relieve vasomotor symptoms compared to placebo, but this did not rank as highly as transdermal oestrogen plus progestogen in the hierarchy of the best treatment options for this outcome. Isoflavones and Black Cohosh were also shown to be more effective when compared to placebo in relief of vasomotor symptoms for women with a uterus, but not significantly better when compared to combined oestrogen plus progestogen. However, the Committee expressed a concern around safety issues of Isoflavones and Black Cohosh and they included a warning about lack of information on their safety profile in the recommendation as well as the large number of different preparations available.

For the outcome of vasomotor symptoms in women without a uterus, although the NMA results were not helpful in decision-making owing to lack of evidence on the most clinically relevant hormonal treatment, which is oestrogen alone (as progesterone is not required for women without a uterus), the direction of effects of non-hormonal treatments was the same for both groups of women (with and without a uterus). The Committee extrapolated the evidence from the network of women with a uterus for their decision-making in selecting the most clinically effective hormonal treatment option for relief of menopausal symptoms in women without a uterus, keeping in mind the limitations on this generalisation. This recommendation was strengthened by the Committee’s clinical experience.

SSRIs or SNRIs were not found to be effective in relieving vasomotor symptoms, but were found to be significantly worse in terms of high discontinuation rates compared to the other treatments for all networks of women (with and without a uterus, and women with a history of breast cancer). Therefore, the Committee decided to draft a recommendation discouraging the use of SSRIs or SNRIs as first-line treatment for relief of vasomotor symptoms for women in menopause. No conclusive points could be made for the outcome of vaginal bleeding for women with a uterus given the limited data for this outcome and the lack of inclusion of several interventions in the network.

For women with breast cancer, there was limited evidence from the NMA assessing the efficacy of different treatments for relief of vasomotor symptoms and the tolerability of these interventions. For the outcome of vasomotor symptoms, only SSRIs/SNRIs, gabapentin, isoflavones and St John’s Wort were included in the network. None of these treatments were found to be significantly better than placebo in relieving vasomotor symptoms for women with a history of, or at high risk of, breast cancer, although St John’s Wort had the highest probability of being the best treatment to achieve this outcome compared to all the other treatments included in the network. However, the Guideline Committee were concerned about the possible interaction of St John’s Wort with other drugs used commonly to treat breast cancer including docetaxel and tamoxifen. Therefore a specific recommendation was
Menopause
Managing short-term symptoms

In relation to treatment of vasomotor symptoms for women with a history or at high risk of breast cancer, the Committee cross-linked to the relevant recommendations in the NICE Guideline on early and locally advanced breast cancer and the NICE Guideline on familial breast cancer. In addition, the Committee discussed thoroughly that women with a history or at high risk of breast cancer should be offered information about all the available treatment options for menopausal symptoms, and referred to a specialist with an interest in menopause for further advice.

In relation to the other short-term outcomes, limited data was found for the outcome of frequency of sexual intercourse, but testosterone was found to significantly increase frequency compared to placebo although the majority of women included in these trials were surgically menopausal. The other evidence identified comparing tibolone versus oestrogen plus progestogen did not show a significant difference in the frequency of sexual activities. Given the limited availability of evidence, the Committee also employed their clinical experience and decided that testosterone should only be offered as an option of improving low sexual desire for women in menopause when HRT is not effective.

Psychological symptoms are common for women in menopause and can impact on their personal, social and professional quality of life. The Committee reviewed the limited evidence available, which showed that low mood can be ameliorated by hormonal replacement therapy (oestrogen alone) and psychological therapies such as CBT but not from the other non-pharmacological treatments reviewed such as herbal treatments. For the outcome of anxiety, psychosocial therapies such as CBT, genistein and red clover were found to significantly reduce anxiety for women in menopause when compared to placebo or usual care. Because of concerns of unknown safety for genistein and red clover, the Committee decided that HRT and CBT are preferable treatment options for low mood for some women in menopause. In addition CBT can be considered for the treatment of anxiety for women in menopause. However, the Committee wanted to draw to the attention of both health professionals and women in menopause that SSRIs/SNRIs should not be a first-line treatment to alleviate low mood for women in menopause who are not diagnosed with clinical depression, owing to their adverse side-effect profile, and cross linked to the relevant NICE Guideline on depression for further advice. The discussion behind this recommendation was that low mood may be the result of hormonal changes occurring at the perimenopause, and antidepressant medication may not be the most appropriate treatment for all women in the absence of clinical depression. However, the group discussed that clinical depression is a different entity and was not assessed as an outcome in this question therefore any conclusion was made only in relation to the outcome of low mood.

The available evidence on efficacy of treatments to relief musculoskeletal symptoms for women in menopause was restricted to one trial that did not show a difference in the experience of these symptoms for women taking oestrogen alone or placebo, independently if they had joint pains at enrolment or not.

Finally, the Guideline Committee expressed concern regarding the wide variety of over-the-counter preparations available and the lack of information on quality control, efficacy, and safety of herbal preparations and complementary therapies. The Committee noted previous reports of fatal adverse events from herbal preparations. Therefore, a separate recommendation was drafted to ensure that women who wish to try complementary therapies are aware that the quality, purity and constituents of these products may be unknown. The Committee discussed that both health professionals and women should be aware of a registration scheme known as the Traditional Herbal Registrations scheme which provides evidence of quality assurance of each of these products in regard to purity rather than the efficacy of a preparation. Preparations registered with the MHRA that have the traditional herbal registrations certificate should be preferred.
In addition, a separate recommendation was drafted in order to raise awareness among women in menopause that biodentical formulations that are compounded for an individual woman according to a health care provider’s prescription are not subject to government regulations or tested for safety or quality and purity of constituents, therefore their efficacy and safety are unknown.

**7.8.3 Consideration of health benefits and resource uses**

There was some evidence suggesting that CBT was more effective than usual care in reducing anxiety and low mood. CBT is not generally available for treatment of low mood in the perimenopausal woman, and the Committee felt that development of psychosocial services such as CBT should be considered given this evidence of benefit. A single CBT appointment in the private sector costs in the region of £80-£90 (http://www.cognitivebehaviouraltherapy.co/fees/) but the GDG thought that it may reduce subsequent contact with health care professionals thus offsetting some of the consultation costs as well as improving health related quality of life. In the absence of any formal economic evidence, the GDG considered that it had the potential to represent a cost-effective use of NHS resources.

Additionally the Committee noted that there are no androgenic preparations specifically licensed for women and that the number of androgenic preparations is decreasing for commercial reasons despite having previously being licensed by the MHRA in the UK.

**Summary of the results of HE model**

The health economic model developed for this guideline suggested that transdermal HRT was the most cost-effective treatment for women with menopausal symptoms with and without an intact uterus, albeit with some uncertainty especially at the lower end of the symptom severity spectrum. However, it should be noted that women with less severe menopausal symptoms are less likely to seek treatment. Transdermal HRT was not found to be cost saving or even 1 of the cheapest treatment alternatives, but it was found to provide the greatest level of benefit in terms of relief of menopausal symptoms for an acceptable additional cost. For women with a uterus transdermal oestradiol and progestogen had an incremental cost-effectiveness ratio (ICER) of £22,165 per QALY relative to the next best non-dominated alternative in the base case analysis and in women without a uterus, transdermal oestradiol had an ICER of £2,149 per QALY relative to the next best treatment option.

A key driver of the results was the network meta-analysis on vasomotor symptoms. This suggested that transdermal HRT was significantly better than placebo. However, the network meta-analysis did not find oral HRT to be significantly better than placebo and this is reflected in the results of the economic model. Nevertheless, there was considerable uncertainty in the network meta-analysis estimates of relative treatment effect and the network meta-analysis did not demonstrate that transdermal HRT was significantly better than oral HR, a cheaper alternative and also the most common first line treatment for relief of menopausal symptoms. Therefore, the GDG did not consider the evidence of the health economic model was sufficiently strong to completely overturn current practice which is reflected in their recommendation that either oral or transdermal HRT can be used.

For women with breast cancer, St John’s Wort appeared to be the most cost-effective alternative treatment, with an ICER of £13,435 relative to the next best non-dominated treatment alternative. This is reflected in a GDG recommendation that women should be advised that St John’s Wort could be considered as a treatment option, whilst highlighting that there is a possibility for it to interact with other medicines.
7.8.4 Quality of evidence

A total of 51 studies were included in the NMA. There were 7 networks constructed for the 3 stratified groups of women in menopause (women with uterus, without uterus and women with breast cancer).

The quality of the NMA was assessed in terms of risk of bias of included trials, heterogeneity of results and inconsistency between direct and indirect evidence. All evidence contributed to the NMA was from randomized trials with a clear description of included population, which was women in menopause excluding those in pre menopause. Data for some treatment comparisons included in the NMA were limited and most of the interventions were compared with placebo and the Committee recognised that this could bias the whole network. In addition, there was a wide variation in the way studies assessed the outcome of vasomotor symptoms, for example, reporting change values in scores, final values or summary measurements such as percentages. The focus of this review question for the NMA was on reporting the frequency of short term symptoms with no inference made to the severity of these outcomes. That was a potential explanation of the increased heterogeneity observed in the networks given the wide variability of baseline characteristics of women in the trials including the wide baseline variation on vasomotor symptoms.

A number of studies on hormonal therapies were not included as they did not meet the inclusion criteria defined in the NMA protocol. However, after scrutinising the rationale of exclusions only a very small minority (2 out of 60) would have been included in a pair-wise meta-analysis and this information was presented to the group as supplementary evidence.

Quality of evidence on pair-wise comparisons for the outcomes of low mood, anxiety, frequency of sexual intercourse and musculoskeletal symptoms was often low or very low due to lack of information on randomisation methods and due to imprecision on estimates of effects. This limited the strength of recommendations that the GDG were able to make for treatments for which only evidence for these outcomes was available.

7.8.5 Other considerations

The recommendations were based on both the interpretation of clinical and health economic analysis of evidence reviewed and on GDG expert opinion.

The Committee was mindful that treatment recommendations should be considered in conjunction with the recommendations on long term outcomes.

The Committee also discussed the potential role of additional information collected for the outcome of VTE, however this was not possible due to lack of reporting of this outcome in short term symptoms studies.

7.8.6 Key conclusions

There is strong evidence that transdermal oestradiol plus progestogen greatly reduces the frequency of hot flushes in women with a uterus. Although there was no strong evidence of efficacy of oral oestrogen plus progestogen treatment, the health economic analysis and the Committee’s expert opinion supported its use in clinical practice. There is also some evidence to suggest that isoflavones and black cohosh may be beneficial for this outcome. The GDG decided to extrapolate the evidence from this population to women without a uterus.

In relation to adverse events, there was also evidence that women with and without a uterus treated with SSRIs had higher rates of discontinuation.

There was relatively limited evidence available for women with breast cancer/history of breast cancer.
For the outcome of low mood, there is some evidence that HRT and CBT may improve low mood for women in menopause.

In relation to anxiety, it was shown that CBT, isoflavones and red clover may improve anxiety for women with menopausal symptoms but there is a lack of consistency between the constituents of herbal preparations, isoflavones and phytoestrogens.

The analysis did not show a difference in musculoskeletal symptoms using oestrogen alone in women who were not recruited because of musculoskeletal symptoms and were older with an average age of 63. Clinical experience suggests that musculoskeletal symptoms may be improved by HRT.

There is evidence that testosterone may increase the frequency of sexual episodes for women in surgical menopause when compared to placebo.

There is very limited evidence for efficacy for SSRIs in symptomatic menopausal women with anxiety but they do not appear to improve low mood in menopausal women who are not clinically depressed.

### 7.9 Recommendations

10. Adapt a woman’s treatment based on her changing symptoms as she goes through the stages of menopause.

11. Offer hormone replacement therapy (HRT) for vasomotor symptoms after discussing the short-term (up to 5 years) and longer-term benefits and risks. Offer a choice of oral or transdermal preparations as follows:
   - oestrogen and progestogen to women with a uterus
   - oestrogen alone to women without a uterus.

12. Do not routinely offer selective serotonin reuptake inhibitors (SSRIs) or serotonin and norepinephrine reuptake inhibitors (SNRIs) as first-line treatment for vasomotor symptoms alone.

13. Explain to women that although there is some evidence that isoflavones or black cohosh may relieve vasomotor symptoms, their safety is unknown and different preparations may vary.


15. Consider cognitive behavioural therapy (CBT) to alleviate low mood and anxiety in menopausal women.

16. Ensure that menopausal women and healthcare professionals involved in their care understand that there is no clear evidence for SSRIs or SNRIs to ease low mood in menopausal women who have not been diagnosed with depression (see the NICE guideline on depression in adults).

17. Consider testosterone supplementation for menopausal women with low sexual desire if HRT alone is not effective.

---

1 At the time of consultation (June 2015), testosterone did not have a UK marketing authorisation for this indication in women. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council’s prescribing guidance: prescribing unlicensed medicines for further information.
18. Explain to women that the efficacy and safety of unregulated compounded bioidentical hormones are unknown.

19. Explain to women who wish to try complementary therapies that the quality, purity and constituents of products may be unknown.

20. Explain to women with breast cancer that St John’s wort may be a treatment option for menopausal symptoms but can interact with other medicines (for example, tamoxifen).

21. For advice on the treatment of menopausal symptoms in women with breast cancer or at high risk of breast cancer, see section 1.13 of the NICE guideline on early and locally advanced breast cancer and section 1.7 of the NICE guideline on familial breast cancer.

22. Offer menopausal women with or at high risk of breast cancer:
   - information on all available treatment options
   - referral to a healthcare professional with expertise in menopause.

7.10 Research recommendations

<table>
<thead>
<tr>
<th>Research question</th>
<th>1. What is the efficacy of different treatments for menopausal symptoms in women who have had treatment for, or are at risk of, breast cancer?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance to 'patients' or the population</td>
<td>Women with a history of breast cancer are currently denied hormonal treatment for menopausal symptoms but the available alternatives are less effective. There is limited evidence from randomised controlled trials on the efficacy of treatments (specifically on vaginal oestrogen) for menopausal symptoms in women who have had treatment for, or are at risk of, breast cancer. There is an urgent need for evidence-based licensed alternatives to traditional HRT in women with breast cancer and other hormone-sensitive malignancies. Randomised controlled trials or large cohort studies are needed to understand the effects of HRT in women with or at risk of breast cancer, and to investigate if there is a difference in breast cancer recurrence, mortality and tumour aggression with different types of HRT.</td>
</tr>
<tr>
<td>Relevance to NICE guidance</td>
<td>High priority: There is an urgent need for evidence-based licensed alternatives to traditional HRT in women with breast cancer and other hormone sensitive malignancies. Research in this area is essential to inform future updates of key recommendations in the guideline.</td>
</tr>
<tr>
<td>Relevance to the NHS</td>
<td>The initial expense of an evidence-based licensed treatment would be offset by reduced visits and hence burden on primary and secondary health care teams and improved workplace productivity.</td>
</tr>
<tr>
<td>National priorities</td>
<td>N/A</td>
</tr>
<tr>
<td>Current evidence base</td>
<td>There is limited evidence from RCTs on the efficacy of treatments (specifically on vaginal oestrogen) for menopausal symptoms in women who have had treatment for, or are at high risk of, breast cancer. In addition, there is a need for a national registry for collecting information on relief of menopausal symptoms and side effects of different treatments used for relief these symptoms in women with breast cancer.</td>
</tr>
<tr>
<td>Equality</td>
<td>Women with hormone-sensitive malignancies are a group in need of special consideration; increasing survival rates should be accompanied by...</td>
</tr>
</tbody>
</table>
### Research question 1.

**What is the efficacy of different treatments for menopausal symptoms in women who have had treatment for, or are at risk of, breast cancer?**

appropriate survivorship management.

#### Feasibility

The proposed research would not require large numbers or duration to demonstrate efficacy. Safety considerations would require a longer trial to demonstrate neutral impact on recurrence rates compared to placebo / no treatment. The main ethical issue is the potential risk of breast cancer recurrence.

#### Other comments

Cancer Research UK could be approached for funding – insufficient funds are currently spent on improving quality of life after breast cancer diagnosis and treatment.

### Research question 2.

**What is the impact of systemic HRT usage in women with a previous diagnosis of breast cancer for the risk of breast cancer reoccurrence, mortality or tumour aggression?**

#### Why this is needed

A number of women with breast cancer experience severe VMS and other menopausal symptoms (usually vaginal atrophy), often due to the treatment they are taking (arimidex/tamoxifen) which greatly reduces their quality of life. Consequently, after trying alternatives unsuccessfully, they opt to use HRT. In those with vaginal atrophy this will be administered locally in the form of an oestrogen cream or pessary, and for other symptoms, e.g. vasomotor symptoms, it will be administered systemically as either combined HRT or a progestogen alone. Most doctors who specialise in the menopause will have a small number of women who opt to receive HRT.

We need to understand the effect of HRT on these women to establish if there is actual increase in
- (1) breast cancer recurrence,
- (2) mortality after breast cancer recurrence,
- (3) tumour aggression.

We need to increase our understanding of route, dosage, at what time-point treatment is initiated, duration of treatment, side-effect profile.

#### Relevance to NICE guidance

High: If current guidance is in fact over-cautious, more women could be offered a treatment that is more effective than those currently available.

#### Relevance to the NHS

The cost for current treatments routinely offered to women with breast cancer such as complementary therapies are frequently borne by women themselves. If HRT were found to be safe in certain women, this increased usage would be offset by the costs of SSRIs, SNRIs, gabapentin, and clonidine which have significant side effects and are used off label.

#### National priorities

N/A

#### Current evidence base

Both cohort studies show no increased risk
2 RCT’s - HABITs trial and Stockholm trial.

#### Equality

Cancer is covered by the Disability Discrimination Act - therefore any women with a breast cancer diagnosis is deemed to have a disability.

#### Feasibility

The risk of increased recurrence or mortality from BC is the major ethical issue.

#### Other comments

Inclusion of women with iatrogenic POI following breast cancer within a registry of women with POI would allow a subgroup analysis.
7.11 **Urogenital atrophy**

7.11.1 **Review question**

What is the effectiveness of local oestrogens and ospemefine compared to placebo for the treatment of urogenital atrophy in women with menopause-related vaginal/urogenital atrophy?

7.11.2 **Introduction to topic**

It is estimated that symptoms caused by vulvovaginal atrophy can affect up to 50% of all postmenopausal women. The most common symptoms affecting the vulva and vagina include dryness, pain on intercourse, vaginal itching and vaginal discharge. There is increased vulnerability to inflammation, trauma and infection. Urogenital atrophy can also result in urinary symptoms such as urgency to urinate and urinary tract infections.

The true incidence and impact of urogenital atrophy on quality of life continues to be underestimated. The reasons for this are believed to be multiple and complex; some women are reluctant to complain about the problem for risk of personal embarrassment, social and cultural reasons, while some healthcare providers are reluctant to bring the problem up in consultation because they are uncomfortable discussing sexual issues.

Symptoms typically become apparent 4 to 5 years after the menopause transition, due to the long-term reduction in oestrogen levels. This results in thinning and loss of elasticity of the vulval and vaginal skin and loss of vaginal lubrication.

There is a range of treatments available, which include local oestrogen therapy (available in creams, pessaries, vaginal ring and tablets) and non-hormonal options such as moisturisers and lubricants. Women may obtain non-hormonal preparations over-the-counter. Treatment should be started early before irreversible changes have occurred and needs to be continued to maintain benefits.

It is vital that sufficient due care and attention is given to this condition to restore and maintain quality of life for increasing numbers of menopausal women in our ageing population.

7.11.3 **Clinical introduction**

The aim of this review question was to assess both the safety and efficacy of local oestrogen treatment and ospemefine for vaginal atrophy. Treatment effects were assessed for treatment duration of less than 1 year (short term) and 1 year or longer (long term).

Outcomes prioritised by the GDG in the short term included patient assessment of symptom improvement, measurement of vaginal pH and maturation index, clinical evaluation of the appearance of the vagina, assessment of endometrial stimulation, breast pain, adverse events, withdrawal from treatment due to adverse events, treatment adherence, treatment acceptability and HRQoL.

Outcomes prioritised by the GDG for long term treatment included improvement in vaginal dryness, dyspareunia and itching and/or discomfort, endometrial thickness, endometrial stimulation, hysteroscopic appearance of the endometrium, endometrial hyperplasia or cancer, withdrawal due to adverse effects, adherence to treatment, acceptability of treatment and HRQoL.

For full details see review protocol in Appendix D.
7.11.4 Description of included studies

7.11.4.1 Short-term effects (less than 1 year)

9 studies in total were included in this review comparing local oestrogens to placebo; 3 RCTs (Dessole 2004, Casper 2009, and Eriksen 1992) included in the systematic review by Suckling 2010 and 6 additional RCTs (Bachmann 2008; Bachmann 2009; Cano 2012; Griesser 2012; Karp 2012; Simon 2008) were identified for inclusion in this review. The systematic review included all randomised comparisons of oestrogenic preparations identified by the search that are administered intra-vaginally for a duration of at least 3 months in postmenopausal women for the treatment of symptoms resulting from vaginal atrophy or vaginitis. The studies included postmenopausal women, who have not menstruated for more than 12 months or who have a serum follicle stimulating hormone (FSH) ≥40IU/l. The definition also included women who have had bilateral oophorectomy (removal of both ovaries). The interventions review included preparations for oestrogen supplementation administered intra-vaginally such as creams, tablets, pessaries and an oestradiol-releasing ring.

The data from the additional 6 RCTs were also added to the meta-analysis. All studies included women with vaginal/urogenital atrophy. The mean age of the women included in these study groups ranged from 56.5 (SD 5.72) years to 66 (SD 7.9) years and the mean time since last menstrual period was between 8.0 (SD 5.8) years to 14.8 (SD 9.6) years but was not reported in 1 study (Griesser 2012) and the range was reported in a second study (Karp 2012) as between 3 and 35 years.

Oestradiol was the most common type of oestrogen preparation used across the studies (Bachman 2009; Bachmann 2009; Simon 2008). Different preparations of local oestrogen were used in the studies including:
- Creams or gel (Bachmann 2009; Cano 2012)
- Vaginal rings (Karp 2012)
- Tablets or pessaries or ovules (Bachmann 2008; Griesser 2012; Simon 2008)

7.11.4.2 Long term effects (more than 1 year)

Two placebo-controlled RCTs were of long-term treatment (52 weeks) (Simunic 2003, Simon 2008) and included in this section.

7.11.4.3 Ospemefine

A total of 8 RCTs comparing ospemefine with placebo were included in this review (Bachmann, 2010; Portman, 2014; Portman, 2013; Rutanen, 2003; Voipio, 2002; Goldstein, 2014; and Simon, 2013, Constantine 2015). Five of these studies (Bachmann, 2010; Portman, 2014; Portman, 2013; Rutanen, 2003; and Voipio, 2002) assessed short-term (< 52 weeks) outcomes of ospemifene treatment; 1 (Simon, 2013) assessed long-term (≥ 52 weeks) outcomes; and two assessed both short- and long-term outcomes (Goldstein, 2014, Constantine 2015). The studies were mainly carried out in United States with 3 studies conducted in Europe (Finland), of which 1 was a multi-site study involving 23 sites. The average age of participants ranged from 40 to 79 years. The majority of studies used 60 mg dosages of ospemifene (Bachmann, 2010; Rutanen, 2003; Voipio, 2002; and Simon, 2013).

7.11.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.
7.11.6 Evidence statements

7.11.6.1 Short term outcomes

Local oestrogens

A meta-analysis of 4 RCTs (of 462 women) found that vaginal pH was significantly reduced in women who received any form of local oestrogen compared with women who received placebo for a treatment period of 12 weeks. The evidence for this finding was of moderate quality.

A meta-analysis of 5 RCTs (of over 600 women) found that the maturation index/value was significantly increased in women who received any form of local oestrogen compared with women who received placebo during a treatment period of 12 weeks. The evidence for this finding was of very low quality.

A meta-analysis of 4 RCTs (450 women) found that the patient assessment of symptom improvement was significantly greater in women who received any form of local oestrogen compared with women who received placebo for a treatment period of 12 weeks. The evidence for this finding was of low quality.

A meta-analysis of two RCTs (of 300 women) found no significant difference in the outcome of endometrial stimulation between women who received local oestrogen and women who received placebo over a treatment period of 12 weeks. The evidence for this finding was of moderate quality.

In terms of adverse events, no significant difference was found for the outcomes of breast pain (moderate quality evidence from 1 RCT of 167 women), adverse events (moderate quality of two RCTs with 321 women), withdrawal due to adverse events (moderate quality evidence from a meta-analysis of 8 RCTs with 1653 women) and treatment adherence (moderate quality of 1 RCT with 43 women) and acceptability (very low quality of meta-analysis of 2 RCTs with 456 women) between women who received local oestrogen and women who received placebo during a treatment period of 12 weeks.

Ospemefine

Pooled analysis of 5 RCTs with 1968 women showed a significant reduction in the percentage of parabasal cells from 60 mg treatment with ospemefine compared to placebo during treatment over a period of less than one year. The evidence was of low quality. The same conclusion was found from 1 RCT with 15 to 16 women which used different doses of ospemefine (evidence was of low to moderate quality).

Pooled analysis of 5 RCTs with 1984 women showed a significant increase in the percentage of superficial cells from 60 mg treatment with ospemefine compared to placebo during treatment over a period of less than one year. The evidence was of very low quality. The same conclusion was found from one RCT with 16 women which used different doses (25, 50, 100, 200 mg) of ospemefine (evidence was of very low quality) although the sample size was too small to allow generalisation of results.

One RCT with 16 women reported a significant reduction in the percentage of intermediate cells at ospemefine doses of 25, 100, and 200 mg compared to placebo during treatment over a period of less than one year, although the effect of 50 mg ospemefine was not significant. Evidence was of very low quality and the sample size of the study was too small to allow generalisation of results.

Pooled analysis of two RCTs with 1149 women showed a significant reduction in the severity of dyspareunia with 60 mg ospemefine compared to placebo during treatment over a period of less than one year. Evidence was of moderate quality.
Pooled analysis of 4 RCTs with 1889 women showed a significant decrease in vaginal pH with 60 mg ospemifene compared to placebo over a treatment period of less than 1 year. Evidence was of moderate quality.

One RCT with 314 women showed a non-significant decrease in the severity of vaginal dryness with 60 mg ospemifene compared to placebo over a treatment period of less than one year. Evidence was of moderate quality.

Six RCTs with 3708 women found significant increases in endometrial thickness associated with ospemifene treatment compared to placebo at dosages of 30, 60, 90, and 200 mg during a treatment period of less than one year. Evidence was of low quality.

No cases of endometrial hyperplasia or carcinoma were reported in any of the 5 RCTs with 1944 women that assessed these outcomes during a treatment period of less than one year. Quality of evidence was of very low quality.

A pooled analysis of 3 RCTs found there was a higher incidence of treatment-related adverse events with ospemifene at 30 mg and 60 mg compared to placebo during a treatment period of less than one year. Quality of evidence ranged from very low to moderate.

Four RCTs found treatment with 30 or 60 mg ospemifene showed no significant increase in the incidence of withdrawal due to adverse events compared to placebo during a treatment period of less than 12 weeks. Evidence was of low quality.

One study with 79 women found no significant differences in Work Ability Index (WAI) for low, anxiety, or self-confidence among ospemifene and placebo groups for the only RCT that assessed this outcome over a treatment period of less than one year. The quality of the evidence was very low.

### 7.11.6.2 Long term outcomes

#### Local oestrogens

A single RCT of over more than 659 women in menopause reported that those using local oestrogens for more than 12 months were significantly more likely to report improvements in vaginal dryness symptoms, dyspareunia and itching or discomfort compared to those on placebo. The quality of the evidence was moderate.

However, this study also found no significant difference for treatment acceptability between women using local oestrogens and those using placebo during 12 months’ treatment. The quality of the evidence was low.

One RCT of 309 women reported no significant difference for the outcomes of endometrial hyperplasia (although one case was confirmed with biopsy at the local oestrogen group) and of treatment withdrawal due to adverse effects associated with the use of local oestrogens compared with placebo for treatment lasted more than 12 months. The quality of the evidence was low.

#### Ospemefine

Pooled analysis of two RCTs with 2560 women showed a significant increase in endometrial thickness associated with 60 mg ospemifene treatment compared to placebo over a treatment period of 1 year. Evidence was of moderate quality.

Two RCTs with over 1000 women assessed endometrial hyperplasia and carcinoma over a treatment period of more than 1 year. One study reported 1 case of endometrial hyperplasia in the ospemifene treatment group. No cases of endometrial carcinoma were reported for each of the two studies. Quality of evidence was low.
Treatment with 30 or 60 mg ospemifene from two RCTs with over 500 women showed no significant increase in the risk of treatment emergent adverse events compared to placebo over a treatment period of more than 1 year. Evidence was of low to moderate quality. The same studies found a significant increase in the risk of withdrawal due to adverse events for the ospemifene group at 60mg compared with placebo over a treatment period of more than 1 year but found treatment with 30 ospemifene showed no significant increases in the incidence of withdrawal due to adverse events compared to placebo when assessed over a treatment period of greater than 1 year. Evidence was of high to low quality. Two RCTs assessed compliance to treatment outcomes over a treatment period of more than 1 year. Compliance was similar in both ospemifene and placebo groups. Evidence was of very low to low quality.

### 7.11.7 Health economics profile

A single search was undertaken for health economic evidence on the treatment of urogenital atrophy in women with menopause-related vaginal/urogenital atrophy. A total of 15 articles were identified by the search. After reviewing titles and abstracts, no papers were obtained. Therefore, no relevant economic evidence was identified for this question.

The following costs for local oestrogens (all are prescription only medications) have been sourced from the British National Formulary (BNF) in November 2014 (Table 16).

### Table 16: BNF costs for local oestrogens

<table>
<thead>
<tr>
<th>Preparation</th>
<th>Price</th>
<th>Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gynest® (Marlborough) Intravaginal cream, estriol 0.01%</td>
<td>Net price 80g with applicator = £4.67</td>
<td>1 applicator full daily, reduced to 1 applicator twice a week</td>
</tr>
<tr>
<td>Ortho-Gynest® (Janssen) Pessaries, estril 500 micrograms</td>
<td>Net price 15 pessaries = £4.73</td>
<td>1 pessary daily, maintenance 1 pessary twice a week</td>
</tr>
<tr>
<td>Ovestin® (MSD) Intravaginal cream, estril 0.1%</td>
<td>Net price 15g with applicator = £4.45</td>
<td>1 applicator-dose daily for 2–3 weeks, then reduce to twice a week</td>
</tr>
<tr>
<td>Vagifem® (Novo Nordisk) Vaginal tablets, f/c, oestradiol 10 micrograms in disposable applicators</td>
<td>Net price 24-applicator pack = £16.72</td>
<td>1 vaginal tablet daily for 2 weeks then reduce to 1 tablet twice weekly; initiate therapy with 10 microgram vaginal tablets, increased after 3 months to 25 microgram vaginal tablet if response inadequate</td>
</tr>
<tr>
<td>Estring® (Pharmacia) Vaginal ring, releasing oestradiol approx. 7.5 micrograms/24 hours</td>
<td>Net price 1-ring pack = £31.42.</td>
<td>worn continuously; replace after 3 months</td>
</tr>
</tbody>
</table>

Local oestrogens are initially used every day for the first two weeks of treatment. The GDG stated that after this time are generally used twice a week. The costs per year for use of each preparation has been calculated and is presented in Table 17.

### Table 17: Cost per year for local oestrogen preparations

<table>
<thead>
<tr>
<th>Preparation</th>
<th>Doses per pack</th>
<th>Cost per dose</th>
<th>Cost per week – used daily for first 2 weeks</th>
<th>Cost per week– used twice a week after first 2 weeks</th>
<th>Cost per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gynest®</td>
<td>16</td>
<td>£0.29</td>
<td>£2.04</td>
<td>£0.58</td>
<td>£33.27</td>
</tr>
<tr>
<td>Ortho-Gynest®</td>
<td>15</td>
<td>£0.32</td>
<td>£2.21</td>
<td>£0.63</td>
<td>£35.95</td>
</tr>
<tr>
<td>Ovestin® (MSD)</td>
<td>3</td>
<td>£1.48</td>
<td>£10.38</td>
<td>£2.97</td>
<td>£169.10</td>
</tr>
<tr>
<td>Vagifem®</td>
<td>24</td>
<td>£0.70</td>
<td>£4.88</td>
<td>£1.39</td>
<td>£79.42</td>
</tr>
<tr>
<td>Estring®)</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>£125.68</td>
</tr>
</tbody>
</table>
7.11.8 Evidence to recommendations

7.11.8.1 Relative value placed on the outcomes considered

Vaginal/urogenital Atrophy is an important cause of reduced quality of life in postmenopausal women of all ages and may require long term treatment. It is due to lack of oestrogen which has significant effects on the vaginal and vulval tissue. Targeted treatment is available with local administration of oestrogen, for example as vaginal cream or tablets. In addition there is now a new agent, ospemifene, which has recently been licensed for use in the USA and Europe (www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/002780/WC500177633.pdf), although it does not currently have marketing authorisation in the UK.

Both short term outcomes (efficacy, safety, acceptability) and long term outcomes (safety, acceptability) of urogenital atrophy were considered in this review question.

In terms of short term symptoms, measurement of vaginal pH, maturation index (for parabasal, intermediate and superficial cells) and women’s subjective assessment of symptoms improvement (relating to atrophy, dryness, dyspareunia (painful intercourse), itching and discomfort were considered as the most important outcomes. Safety outcomes included an assessment of endometrial stimulation, breast pain as a surrogate marker for systemic absorption, and blood oestradiol levels, adverse events (including withdrawal due to adverse events), acceptability and adherence to treatment. In terms of long term outcomes, endometrial hyperplasia or cancer, long term relief of symptoms and impact on the health quality of life were considered.

7.11.8.2 Consideration of clinical benefits and harms

The included evidence showed that local vaginal oestrogen was beneficial for improving short term outcomes (vaginal pH, maturation index, patients’ symptomatic improvement) for menopausal women when compared to placebo. Furthermore, no difference in the experience of adverse events was found between those women treated with local vaginal oestrogen and those on placebo. In terms of long term outcomes, although a significant improvement was found for women who used local vaginal oestrogens compared to placebo groups in relieving vaginal dryness symptoms, dyspareunia and itching or discomfort, there was also a case of endometrial hyperplasia (although the difference in this outcome was not significant) for those who used local oestrogen treatment compared to placebo. Endometrial hyperplasia is an abnormal proliferation of endometrium and is considered as a risk factor of endometrial cancer.

The GDG concluded that, given the effectiveness of vaginal local oestrogen in relieving symptoms of urogenital atrophy and the reasonable safety profile, it should be considered as a treatment for this condition for women in menopause (including those who take systemic HRT but experience persistent urogenital symptoms).

The GDG recognised the need to inform women that treatment for urogenital atrophy with local oestrogen does not provide permanent relief from this symptom, which may recur after discontinuation of local oestrogen treatment. The GDG discussed duration of treatment and concluded that local oestrogen can be used in the long term to provide symptom relief; since systemic absorption of oestrogen from recommended doses is very small, it is unlikely to be associated with the adverse effects reported with the use of systemic HRT. Although there was some evidence that showed one case of endometrial hyperplasia with the long-term use of local oestrogens, this risk was too low to suggest regular monitoring of all women using local oestrogens and the GDG considered in their expert opinion that ultrasound measurement of endometrial thickness was not necessary during treatment with vaginal oestrogen. This is true of all local oestrogen preparations as several are currently available. With regard to adverse events, the group wished to inform women who may opt for this local
treatment that adverse events are considered rare, but as is the case with systemic HRT, unscheduled vaginal bleeding should be reported to a health care professional. A separate recommendation was drafted to highlight this guidance.

The group also discussed the dose and administration of vaginal local oestrogens, and discussed management if symptoms of urogenital atrophy still persist despite standard treatment with local oestrogen. The GDG considered that it would be a safe option to increase the dose, noting that vaginal tablets of oestradiol are currently available only at a dose of 10 micrograms and that a dose of 20 micrograms may be required, as supported by the reviewed evidence. However the group concluded that the most appropriate way to manage persistent symptoms would be to obtain advice from a health professional with expertise on this area.

The group discussed the role of local oestrogens for women in whom systemic HRT is contraindicated, for example women with a history of breast cancer, and concluded that low dose local oestrogen should still be considered for relieving symptoms of urogenital atrophy in these women, as there is minimal systemic absorption of low dose preparations, although this decision should be discussed with a health professional with expertise in the field as even very small amounts of oestradiol may decrease the effect of aromatase inhibitors which are used in the treatment of breast cancer.

Moisturisers and lubricants were considered a safe option for this condition by the GDG, The GDG discussed also the place of ospemifene in the treatment of urogenital atrophy. The group concluded that ospemifene could be considered in the context of being an oral therapy but with associated side effects similar to those of systemic HRT preparations rather than to those of local preparations. Although ospemifene is a tissue-selective oestrogen receptor agonist/antagonist that was found to exert a beneficial effect on relieving symptoms of urogenital atrophy through significant decrease of parabasal and superficial cells and by lowering the vaginal pH, it was also found to significantly increase the endometrial thickness and was associated with higher risk of withdrawals due to adverse events compared to the placebo group. Although different dosages of ospemifene were included, the majority of evidence was based on the dosage of 60 mg per day The GDG did not consider recommending ospemifene for the treatment of urogenital atrophy, given the overall balance of clinical benefit and risk from the associated adverse events, and noted that ospemifene currently does not have UK marketing status.

7.11.8.3 Consideration of health benefits and resource uses

The costs for local oestrogens can be low, ranging from approximately £33 to £170 per year and symptoms of urogenital atrophy will lead to a reduction in women’s HRQoL. Local oestrogen has been shown to have health benefits with a low risk of adverse events, therefore the GDG considered that recommending these treatments is likely to be a cost-effective use of resources.

A price for ospemifene could not be identified as it is not currently available in the UK. However, the GDG understand that where it is available it is at a high cost. The health benefits and resource use associated with ospemifene need to be compared to the other local oestrogens available for treating vaginal atrophy. Potential serious side effects of a systemic treatment will need to be balanced against those associated with treatment administered locally.

7.11.8.4 Quality of evidence

The majority of evidence was of moderate to very low quality. Different dosages of local oestrogens were used but data were too limited to allow further subgroup analyses. There was a paucity of information for any long term follow-up data longer than 1 year of treatment, therefore the results of the included studies should be interpreted with caution given the
unknown long term efficacy and safety of this treatment. Detection bias, inconsistency and
imprecision were the main domains downgraded in GRADE quality assessment.

7.11.8.5 Other considerations
The recommendations were based on both the interpretation of clinical evidence reviewed
and on GDG expert opinion.

The GDG wished to emphasise the importance for health professionals of discussing
urogenital atrophy with women in menopause, despite a reluctance to bring up sexual issues
in consultation, as women may be unaware that there are effective treatments. Patient
preferences should be considered in advising women on methods of administration of local
vaginal oestrogen.

The group discussed that, when vaginal oestradiol tablets are recommended, 10mcgs is now
the routine dose preparation available. This dose may need to be increased after specialist
advice following failure of the lower dose to relieve the symptoms of urogenital atrophy.

7.11.8.6 Key conclusions
The group concluded that vaginal local oestrogens were found effective in relieving
symptoms in the short term and long term (up to a year) for women in menopause with
urogenital atrophy without risking the safety outcomes for this population. The GDG
concluded that although ospemifene may be beneficial in terms of improving outcomes for
short term urogenital symptoms, adverse events had to be considered in the balance of
clinical benefit and risk, and the non UK marketing status precluded the GDG from
formulating any recommendations.

7.11.9 Recommendations

23. Offer low-dose vaginal oestrogen to women with urogenital atrophy (including
those on systemic HRT) and continue treatment for as long as needed to relieve
symptoms.

24. If systemic HRT is contraindicated, consider low-dose vaginal oestrogen after
seeking advice from a healthcare professional with expertise in menopause.

25. If low-dose vaginal oestrogen does not relieve symptoms of urogenital atrophy,
consider increasing the dose after seeking advice from a healthcare professional
with expertise in menopause.

26. Explain to women with urogenital atrophy that:
   - symptoms often come back when treatment is stopped
   - adverse effects from low-dose vaginal oestrogen are very rare
   - they should report unscheduled vaginal bleeding to their GP.

27. Advise women with vaginal dryness that moisturisers and lubricants can be used
alone or in addition to vaginal oestrogen.

28. Do not offer routine monitoring of endometrial thickness during treatment for
urogenital atrophy.
8 Review and referral

8.1 Review question
At what intervals should clinical review be undertaken to assess the effectiveness and safety of treatments to relieve menopausal symptoms and to determine when women need to be referred to specialist care?

8.2 Introduction to topic
It is important for women to be involved in the ongoing management of menopause symptoms, as partnership with their health professionals will maximise health benefits, improve compliance with medication and address any adverse effects appropriately.

Following the decision to start treatment for menopause symptoms, initial review is required to assess efficacy and side effects. With hormonal therapy, unscheduled bleeding is common in the first 3 months and adjustment of dosage or formulation may be considered if there are persistent side effects such as bloating, nausea and breast discomfort. Once treatment is established, review is necessary to assess changes in risk due to new or pre-existing health problems, to carry out basic health checks such as measurement of weight and blood pressure and to inform and engage women in national screening programmes. If hormonal therapy was initiated in the perimenopause, change from cyclical to continuous combined HRT should be discussed at review and with longer duration of therapy a reduction in the dosage of oestrogen in the HRT may be considered. Information provision and discussion of longer term risks and benefits is important. Regular review of the efficacy and safety of non-hormonal treatment is also required.

There is also a need to provide women with information about when and how professional help should be accessed if problems such as unscheduled bleeding or other risk factors or concerns arise in between scheduled visits.

8.3 Clinical Introduction
The objective of this review is to determine at what intervals should clinical review should be undertaken to assess the effectiveness and safety of treatments and when women need to be referred to specialist care. The search for this question included RCTs, comparative prospective or retrospective studies.

For full details see review protocol in Appendix D.

8.4 Description of included studies
No studies met the inclusion criteria for this review and no evidence table was generated.

8.5 Evidence statements
No studies were identified for this review question and so there is no evidence profile.
8.6 Evidence to recommendations

8.6.1 Relative value placed on the outcomes considered

The GDG selected the following outcomes for this question; number of unscheduled hospital appointments, continuation with treatment, health quality of life and adverse events.

8.6.2 Consideration of clinical benefits and harms

Given the absence of clinical evidence for this review topic, the GDG discussed the importance and timing of review of treatment for short term menopause-related symptoms. The GDG were aware of guidance on the follow-up of women taking HRT, including advice on specialist referral, provided by specialist organisations.

The GDG considered that 3-month review after initiation of treatment is currently the routine review practice. The timing of this review was supported by the GDG, in order to assess the clinical effectiveness and tolerability of treatments and act on alternative treatment options if treatment was unsuccessful or not well-tolerated.

The group considered ongoing review and concluded that women continuing on treatment for menopausal symptoms should be reviewed annually, unless there are clinical indications for earlier review. The GDG discussion concluded that women should be referred to a specialist if they experience a serious adverse effect, and should also be referred in the event of persistent treatment failure, or if they experience ongoing adverse events.

At the review visit, for those women still symptomatic for menopausal symptoms and who are medically unsuitable for HRT, the GDG considered that information on alternative treatments for the relief of menopausal symptoms should be offered and referral should be considered to healthcare professional with experience in menopause.

8.6.3 Consideration of economic benefits and harms

The frequency of clinical review to assess the effectiveness and safety of treatments to relieve menopausal symptoms and to determine when women need to be referred to specialist care clearly has implications for health care resources. If the frequency of clinical review is too great then additional resources will be used for insufficient gain. Conversely, if the frequency is insufficient then the patient may continue on ineffective treatment longer than necessary. However, in the absence of clinical evidence then it is difficult to suggest an optimal frequency and the GDG relied extensively on their clinical experience, expert opinion and existing guidance in order to make recommendations. The GDG did not think that their recommendations would represent a change in the current practice and therefore would not result in additional resource implications for the NHS.

8.6.4 Quality of evidence

No clinical evidence was found for this review question.

8.6.5 Other considerations

The recommendations were based on GDG expert opinion.

The group also considered the importance of advising women about the value of recommended health screening including national programmes such as the NHS Breast screening programme that is offered every 3 years to women aged 50 years or over and the Cervical screening which is available for all women aged 25 years and over with a frequency of routine 3-yearly recall between ages 25-49, then 5-yearly recall until aged 65.
8.6.6 Key conclusions

The GDG concluded that in the absence of relevant evidence, recommendations were based on GDG’s clinical experience, expert opinion and existing guidance.

8.7 Recommendations

29. Discuss with women the importance of keeping up to date with nationally recommended health screening.

30. Review each treatment for short-term menopausal symptoms:
   • at 3 months to assess efficacy and tolerability
   • annually thereafter unless there are clinical indications for an earlier review (such as treatment ineffectiveness, side effects or adverse events).

31. Refer women to a healthcare professional with expertise in menopause if treatments do not improve their menopausal symptoms or they have ongoing troublesome side effects.

32. For women with menopausal symptoms and contraindications to HRT:
   • provide information on non-hormonal and non-pharmaceutical treatments (for example, CBT, hypnosis, acupuncture and relaxation techniques) for the relief of menopausal symptoms
   • consider referral to a healthcare professional with expertise in menopause.

33. Consider referring women to a healthcare professional with expertise in menopause if there is uncertainty about the most suitable treatment options for their menopausal symptoms.
9 Starting and stopping HRT

9.1 Review question

In perimenopausal and postmenopausal women using Hormonal Replacement Therapy (HRT) for vasomotor symptom relief, what is the effectiveness of an abrupt HRT discontinuation strategy compared with a tapered HRT discontinuation strategy?

9.2 Introduction to topic

The majority of women taking hormone replacement therapy will decide, at some stage, to discontinue their treatment. When a woman wishes to continue, there should be a discussion about the benefits and risks for that individual woman and what she could expect if she stops treatment. In recent years, many women were advised to stop HRT after 2-5 years of use or at the age of 60 although the evidence for this advice is uncertain.

The options for discontinuation are either for the women to stop treatment immediately or to gradually wean off treatment by decreasing the dose or number of days per week that HRT is taken. The discontinuation process can range from several weeks to several months.

9.3 Clinical introduction

The aim of this review was to determine the optimal method of stopping hormone replacement therapy in women who have been using HRT for menopause symptom relief. Specifically, the review aimed to determine whether abrupt discontinuation of HRT and tapering the dose of HRT (by any method) were more effective.

RCTs and comparative cohort studies available were considered appropriate study designs to address this review. The GDG considered the following outcomes to be critical: recurrence of menopausal symptoms, HRQoL, re-uptake of HRT or use of alternative treatment and acceptability of the interventions to women.

For full details see review protocol in Appendix D.

9.4 Description of included studies

Four RCTs comparing abrupt discontinuation with tapered discontinuation were included in the review. The studies were conducted in Brazil (Chuha 2010), Israel (Haimov-Kochman 2006), Sweden (Lindh-Åstrand 2010), and Turkey (Aslan 2007). No prospective cohort studies were found that met this protocol.

We found evidence for the outcomes of recurrence of menopausal symptoms, recommencement of HRT, and the impact on women’s HRQoL. No evidence was found on the uptake of alternative treatments or acceptability of the different discontinuation methods to women.

Although all included RCTs compared abrupt HRT discontinuation to a tapering dose, the method of tapering and the follow-up time of outcomes differed widely. The characteristics of the included studies are thus reported separately.

More specifically, the majority of trials (3 out of 4) (Haimov-Kochman 2006, Aslan 2007, Cunha 2010) reported results on menopausal symptoms both during the tapering process and at the end of this process (when both groups had stopped HRT) whereas 1 study (Lindh-Åstrand 2010) reported results when HRT discontinuation was completed in both groups.
One trial (Haimov-Kochman 2006) included 91 women (50 were randomised to abrupt discontinuation and 41 to dose tapering) with a mean duration of HRT use of 8.8 years (SD 3.8). Approximately 50% of women had commenced HRT because of symptomatic hot flushes. The method of tapering used in this study involved reducing the HRT dose by 1 tablet per week each month, and discontinuation was completed over a 6 month period; women were followed-up for a total of 12 months. Occurrence of menopausal symptoms was assessed at 1, 3 and 6 months (during the tapering process) and again at 9 and 12 months (when both groups had stopped HRT completely).

Another trial (Aslan 2007) (total sample size of 70 women equally randomised to two HRT discontinuation methods) considered women taking a tablet every other day for two weeks, prior to stopping treatment altogether. The mean duration of HRT use was 6.3 years (SD 0.68) in women who immediately discontinued HRT, and 5 years (SD 0.52) in those who were randomised to tapered discontinuation. 79% of women reported suffering with vasomotor symptoms prior to commencing HRT. Follow-up of this study was for 4 weeks. Occurrence of menopausal symptoms was assessed at 2 weeks (during tapering) and at 4 weeks (when both groups had finished treatment).

The trial by Cunha 2010 (total sample size of 60 women equally randomised to two HRT discontinuation methods) compared 3 methods of discontinuation: immediate discontinuation of standard dose HRT, conversion to low dose HRT for 2 months prior to discontinuation, and conversion to low dose HRT for 4 months prior to discontinuation. The mean duration of HRT use was 4.8 years, with no significant differences between the 3 groups. An inclusion criterion for this study was that HRT had been prescribed for the treatment of climacteric vasomotor symptoms. Follow-up was for 6 months. Occurrence of menopausal symptoms was reported during tapering (at 2 months and 4 months) as well as after complete discontinuation at 6 months.

The last study by Lindh-Åstrand 2010 compared immediate discontinuation of HRT with dose tapering (taking 1 tablet every other day for 4 weeks before stopping). 87 women participated in this trial; 41 were randomised to immediate discontinuation and 46 were randomised to the taper-down strategy. Mean duration of HRT use was 9.2 years, with no significant differences between the two groups. As with the previous study, an inclusion criterion was that HRT had originally been prescribed because of vasomotor symptoms. Follow-up was conducted at 6 weeks after discontinuing HRT completely (i.e. 6 weeks after randomisation for the immediate discontinuation group, and 6 weeks after completion of tapering in the taper-down group). Therefore no data on symptom severity during the tapering process were available.

9.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F; study evidence tables in Appendix H; forest plots in Appendix J; and exclusion list in Appendix G.

Data comparing the two methods during tapering the HRT discontinuation, and following tapering are presented in two GRADE tables, with a brief description of the tapering duration (and time point at which the measurement was taken) provided for each outcome.

9.6 Economic evidence

No health economic search was undertaken for this question.
### 9.7 Evidence statements

#### Menopausal symptoms during tapering process

Low quality evidence from a RCT with a small sample size (35 participants) showed no significant difference in overall scores of a scale of menopausal symptoms (Blatt Kupperman Index) at 2 or 4 months for those who were on tapered HRT when compared with those who had stopped HRT abruptly. However, the same study showed a significant decrease in scores for the vasomotor component of the Blatt Kupperman Index at both 2 and 4 months for women on tapered HRT compared with those who stopped HRT abruptly (moderate quality evidence).

Low quality evidence from another RCT including 70 women showed no significant difference in the overall number of hot flushes during 2 weeks when comparing the tapered and abrupt discontinuation groups. The same study also found no significant difference between the two groups in the number of women with either no symptoms, mild, moderate or severe vasomotor symptoms at 2 weeks (very low quality evidence).

Low quality evidence from a RCT enrolling 91 participants showed a significant reduction in the total and vasomotor component of the Greene Climacteric Score at 1 month and 3 months for women who were tapering HRT over 6 months, as compared to those who had stopped abruptly. However, low quality evidence from the same study showed that, at 6-month follow-up, while there was no significant difference between the two groups in total scores of the Greene Climacteric Scale, a significant increase in scores for the vasomotor component of the scale was found in the tapered HRT group compared with the abrupt discontinuation group. Results should be interpreted with caution as only narrative summaries of these results were provided.

#### Menopausal symptoms after discontinuation

Low to very low quality evidence from the same RCTs that provided evidence during discontinuation (see evidence statements above) found no significant difference between groups for any of the outcomes (scales of total vasomotor symptoms; hot flushes; proportion of women with none; mild or severe vasomotor symptoms) for between 4 weeks and 12 months after discontinuation.

Low quality evidence from one RCT with 81 participants showed no significant difference in the impact of women’s HRQoL at 6 weeks when treatment was tapered over 4 weeks compared with a strategy of abrupt discontinuation. Results for outcomes at 6 weeks, 9 and 12 months and health quality of life should be interpreted with caution as only narrative summaries of the results (with no measures of effect size) were provided in the published papers.

Low quality evidence from two RCTs including 171 participants showed no significant difference between groups in the number of women who recommenced HRT treatment at 12 months after HRT had been discontinued either by tapered (over 4 weeks or 6 months) or abrupt methods.

### 9.8 Evidence to recommendations

#### 9.8.1 Relative value placed on the outcomes considered

The GDG selected the following outcomes as the most important for their decision making: reoccurrence of menopausal symptoms which may or may not result to in resumption of HRT treatment, uptake of alternative treatment, acceptability of method of discontinuation of HRT treatment, by women and the impact on women’s their HRQoL.
9.8.2 Consideration of clinical benefits and harms

The GDG initially discussed the importance of informing women about the occurrence of unscheduled vaginal bleeding as a common side-effect of HRT use within the first three months of treatment. They decided that this should not be a reason for HRT discontinuation, although it should be reported at the initial review with the health care professional (3 months).

There was some evidence from randomised participants which suggested that the method of tapering the HRT treatment until discontinuation made no difference to a woman’s experience of total menopausal symptoms (including vasomotor symptoms) in the long term (at around 6 months) when compared to abrupt discontinuation. In addition, 1 other study found that there was no significant difference in the proportion of women with either none, mild or severe menopausal symptoms in the tapering compared with the abrupt discontinuation treatment groups. However, it was found that, specifically for vasomotor symptoms, there may be some improvement in the relief of menopausal symptoms in the shorter term (between 2 to 4 weeks after stopping HRT abruptly) and longer term (6 months) associated with the tapering compared to abrupt HRT discontinuation method.

In terms of the reoccurrence of menopausal symptoms and the impact on women’s HRQoL, none of these symptoms seem to reappear impact after the end of HRT discontinuation using either method. Furthermore, no adverse effects were found to be associated with either method, although the evidence was scarce.

The GDG discussed the impact of each method of HRT discontinuation on women’s symptoms and on other outcomes but given that no strong evidence for an improvement in outcomes or harm was found to be associated with one method or another, they emphasised the importance of a woman’s personal preferences and consider both methods of HRT discontinuation (abrupt and tapering) should therefore be considered. In addition, they didn’t feel that identifying a specific duration for the tapering method of HRT discontinuation would be informative given that this will vary depending on women’s personal circumstances and their needs and tapering methods may involve either cutting the dose or administering HRT less often and may take weeks to months for complete discontinuation.

9.8.3 Consideration of economic benefits and harms

Compared to immediately stopping treatment, a tapering regimen will necessitate the taking of hormone replacement of HRT for longer and therefore require a greater treatment cost. However, in the absence of good quality evidence the GDG were of the view that a tapered approach could result in a lower recurrence of symptoms in the short term which could potentially reduce other health care use to offset the treatment costs and result in a better HRQoL. The GDG therefore thought that either approach could be offered to the woman.

9.8.4 Quality of evidence

The quality of the evidence supporting these recommendations was generally of low to very low quality given that there was a high risk of bias (most studies were unblinded) and lack of precision in the estimates of effects. In addition, most of the results found the outcomes reported after HRT discontinuation did not provide measures of variation of effects. Because these outcomes were only reported in narrative summaries and that in some cases, the level of statistical significance was the only data presented, these results should be interpreted with caution. There was also variation of the timing of outcomes reported depending on different patterns of tapering of HRT treatment.
9.8.5 Other considerations

The recommendations were based on both the clinical evidence reviewed and GDG expert opinion.

The GDG discussed that because no further analysis was conducted on the type of HRT taken by women prior to decision of HRT withdrawal, no conclusions are made based on this factor. The GDG discussed that there may be some differences in these two methods of HRT discontinuation (tapering, abrupt) based on the previous type of HRT use (for example, if women have previously used stronger oestrogen replacement therapies) and on HRT duration and this needs to be taken into consideration.

9.8.6 Key conclusions

The GDG concluded that the evidence was not conclusive and the GDG used clinical experience to inform their decision making.

9.9 Recommendations

34. Explain to women with a uterus that unscheduled vaginal bleeding is a common side effect of HRT within the first 3 months of treatment but should be reported at review appointments.

35. Offer women who are stopping HRT a choice of gradually reducing or immediately stopping treatment.

36. Explain to women that:
   - gradually reducing or immediately stopping HRT makes no difference to their symptoms in the longer term
   - gradually reducing HRT may limit recurrence of symptoms in the short term.
10 Long-term benefits and risks of HRT

10.1 Venous thromboembolism

10.1.1 Review question

What are the effects of HRT administered for menopausal symptoms on the risk of developing venous thromboembolism (VTE)?

10.1.2 Introduction to topic

VTE is a condition comprising deep vein thrombosis (DVT) and pulmonary embolism (PE), precipitated by conditions that cause blood flow to slow down such as immobility, compression of the blood vessel or increased blood viscosity. This causes vascular endothelial damage, coagulation of the blood and clot formation, the latter sometimes breaking up resulting in clots lodging in the lungs alterations in the constituents of the blood (i.e., inherited or acquired alterations in coagulation) also increase the chance of clot formation. The clotting mechanism is designed to stem haemorrhage from damaged vessels and functions as a fine balance between the clotting cascade, and the fibrinolytic system which acts to counter balance this ensuring the clot remains localised and does not spread to obstruct the entire vessel. Alterations of this delicate homeostatic balance can be both inherited and acquired.

The most frequent causes of an inherited thrombophilia are known gene mutations of factor V Leiden and prothrombin, which together account for 50% of cases. Defects in the natural anticoagulants protein S, protein C, and antithrombin III deficiency account for most of the remaining cases, with rare disorders of fibrinogen. There may be additional causes.

Risk factors for thrombosis include increasing age, surgery, trauma, immobilisation, malignancy, pregnancy, hormone use, obesity, smoking, antiphospholipid syndrome, and a number of other major medical illnesses.

Studies that have evaluated the association between HRT and VTE have suggested that HRT caused an approximately twofold increase in VTE risk, which appeared to be greatest in the first year of treatment (Canonico 2008) as well as increased BMI (Canonico 2006).

Oral HRT is ingested and metabolised in the liver. Whilst undergoing first pass metabolism it affects the clotting cascade by increasing resistance to protein S and protein C (natural anticoagulants) and increasing fibrinogen, thus increasing a woman’s tendency towards thrombosis. Transdermal oestrogens are absorbed directly into the blood stream, thus avoiding this first pass metabolism and therefore having less effect on the coagulation factors in the liver.

In addition to the type and route of oestrogen administration, the type of progestogen may also affect the risk of VTE (Canonico 2006).

10.1.3 Clinical introduction

The aim of this review was to determine the effect of HRT on the risk of developing VTE for women in menopause. The focus population of the review was women who have initiated HRT use before the age of 65 years. Given that the risk of developing VTE may be different for women of different ages or at different menopausal stages, the GDG decided at the protocol stage to look at subgroup analyses on age groups and user categories (i.e. ever, past and current users). The risk of developing VTE was examined in terms of different HRT types, duration and timing since discontinuation.
RCTs and comparative prospective cohort studies were selected for inclusion in this review.

With regards to selection of comparative cohort studies, only those that have adjusted for the most common confounders for developing VTE such as age, BMI, family history of VTE in their analyses were included.

Two outcomes were prioritised by the GDG: risk of developing VTE (including DVT and PE) and mortality related to VTE.

For full details see review protocol in Appendix D.

### 10.1.4 Description of included studies

Seven RCTs comparing some form of HRT with placebo were included in this review.

All included RCTs assessed the effect of HRT in comparison with placebo group (Cherry 2002; Høibraaten 2000; Holmberg 2008; Manson 2013 (the WHI, including 6 intervention or post-intervention reports: Anderson 2004, Canonico 2014, Curb 2006, Cushman 2004, Heiss 2008, LaCroix 2011); Nachtigall 1979; Vickers 2007; Whiteman 1999)

Two RCTs (Cherry 2002; Manson 2013) compared the effect of oestrogen alone versus placebo.

Four RCTs (Høibraaten 2000, Nachtigall 1979, Manson 2013, Vickers 2007) compared oestrogen plus progestogen in comparison with placebo.

The majority of RCTs were conducted in the USA: (Manson 2013, Nachtigall 1979, Whiteman 1999), and others in the UK (Cherry 2002), Norway (Høibraaten 2000), and Sweden (Holmberg 2008). One was a multi-centre study carried out in the UK, Australia and New Zealand (Vickers 2007).

Eight comparative cohort studies (Benson 2012, Canonico 2009, Eischer 2014, Grodstein 1996, Laliberté 2011, Ohira 2010, Olié 2011, Su 2012) were included which compared HRT with no treatment were included. They were conducted in the USA, France, UK, Austria, Canada, and Taiwan. Sample sizes of the included cohort studies varied widely, ranging from 630 (Eischer 2014) to 105,825 (Benson 2012, Million Women Study).

The majority of included studies considered first incidence of VTE in HRT users. Three studies examined the risk of VTE in women at a higher risk for VTE, for example those with a previous history of VTE (Høibraaten 2000 (RCT), Eischer 2014 (prospective cohort study)), and breast cancer (Holmberg 2008, RCT). Most of the included studies have excluded pre- or perimenopause women.

When studies have provided results by HRT type (oestrogen alone or oestrogen plus progestogen) and the associated risk on VTE for women in menopause then these are presented separately.

### 10.1.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.

### 10.1.6 Economic evidence

No health economic search was undertaken for this guideline as the decision was made to prioritise short term treatment. However, the clinical evidence from this review was used to inform the model on short term treatment where there was impact on longer term outcomes arising from short term use.
Menopause
Long-term benefits and risks of HRT

10.1.7 Evidence statements

10.1.7.1 Evidence statements for RCTs

Current use of oral HRT

Moderate to very low quality of evidence from 7 RCTs including 34,379 women showed a
significantly increased risk of VTE with current oral use of any HRT when compared with
placebo. The same result was found when the role of either oestrogen alone or oestrogen
plus progestogen was examined in comparison with placebo by two RCTs (including 11,756
women), and 4 RCTs (including 21,301 women), respectively.

Duration of HRT use

Findings on the risk of VTE in relation to duration of HRT use were mixed. Moderate quality
evidence from a single RCT including 4,385 participants showed a significantly increased risk
for up to 1 year duration, and more than 5 years duration (evidence from 2 other RCTs
including more than 20,000 participants). However, low quality evidence from 4 RCTs (with
2,479 participants) showed no significant difference between those who were on HRT
between 1 and 5 years when compared with those on placebo.

Women aged 50-59 years

When the subgroup of women aged 50-59 years at baseline was examined, low quality
evidence based on over 5000 women from a RCT showed an increased risk in VTE for
women taking oestrogen plus progestogen in comparison with those in placebo group,
whereas findings based on over 3000 women from another RCT showed no significant
difference on the VTE risk between oestrogen alone use and placebo (very low quality
evidence).

Time since menopause

One RCT with a subgroup analysis on women at 50-59 years showed that among those who
have initiated oestrogen plus progestogen within 10 years since menopause, there was a
significant increased risk of VTE when compared with the placebo group (moderate quality
evidence). However, very low quality evidence from the same RCT showed that the risk of
VTE was not significantly different between oestrogen alone users and placebo groups.

10.1.7.2 Evidence statements for comparative cohort studies

Current HRT use

Moderate to very low quality evidence from 3 cohort studies with sample sizes ranging from
over 600 to almost 800 and 60,000 women with menopause found a significantly increased
risk of VTE in current HRT users in comparison with no treatment group.

Past use of HRT

Moderate to very low quality evidence from 5 cohort studies (sample sizes ranged from 6,600
to half a million) all showed no significant difference in the risk of VTE for past users of HRT
(no information on the type of HRT administration), as compared to never users.

Administration routes of HRT

Low quality evidence from two cohort studies found a significantly increased risk of VTE
among oral HRT users in comparison with non-users, whereas low to very low quality
evidence from 3 cohort studies with sample sizes up to more than half a million women showed no significant difference between transdermal HRT users and nonusers. When oral and transdermal HRT uses were compared head-to-head in two cohort studies including more than 54,000 participants, both studies showed a significantly increased risk of VTE in oral HRT use in comparison with transdermal use (low quality evidence).

Subgroup analysis from 1 cohort of more than half a million women by age distribution (women aged less than 50 years or over 50 years) showed similar results regarding the effect on VTE from oral and transdermal HRT use; more specifically, oral HRT use was found to increase the risk of VTE when compared to no treatment, whereas this was not the case for the comparison of transdermal HRT use and no treatment (moderate to low quality evidence).

**Types of HRT**

Low to very low quality evidence from two cohort studies including more than half a million of women found a significantly increased risk of VTE among oestrogen users in comparison with non-users, whereas another 3 cohorts showed no significant difference between oestrogen users and nonusers (very low quality evidence). However, the results from the 3 last cohorts that found no significant difference in the risk of VTE for oestrogen users and no users should be interpreted with caution as they have used both types of HRT administration (transdermal and oral) and the PE was included in their analysis. One study also presented results for the outcome of DVT and found a significantly increased risk of VTE among oestrogen users compared with non-users (moderate quality evidence).

When oestrogen plus progestogen was examined, low quality evidence from 1 cohort study (with more than 670,000 women) showed a significantly increased risk of VTE in oestrogen plus progestogen users in comparison with nonusers. However, another cohort study carried out among Chinese women found no significant difference between oestrogen plus progestogen users and non-users (very low quality evidence).

**Duration of HRT**

Moderate quality evidence from a large cohort study (about half a million women) showed a significantly increased risk of VTE in oral HRT users with a duration of 2 years or less compared to non-users, but this difference was not found significant when transdermal HRT users were compared with non-users. Similar effect direction was found when different administration routes (oral and transdermal routes) were analysed in terms of HRT duration up to 5 years and more than 5 years compared to non-users. This evidence was of moderate to very low quality based on two cohort studies.

Subgroup analysis from 1 study with more than half a million women cohort on the age distribution (women aged less than 50 years or over 50 years) showed similar results regarding the effect on VTE from oral and transdermal HRT use; more specifically, oral HRT use was found to increase the risk of VTE when compared to no treatment arm, whereas this was not the case for the comparison of transdermal HRT use and no treatment (moderate to low quality evidence)

**Recurrence of VTE among women who have had a first VTE**

Very low quality evidence from a single cohort study including 630 women in menopause found no significant difference in the risk of recurrence of VTE among oestrogen alone users who have had a first VTE compared with non-users.
Different preparations of oestrogen and progestogen in combined HRT

When different preparations of oestrogen were examined, low quality evidence from a cohort study of over half million women showed a significant increase in the risk of VTE for users of conjugated equine oestrogens and oestradiol users compared to no treatment (low quality evidence).

With regards to different types of progestogen in combined HRT (oestrogen plus progestogen), mixed findings were reported across studies:

- Very low quality evidence from two cohort studies showed no significant difference in the risk of VTE for current users of micronised progesterone as a component of combined HRT in comparison with non-users.
- The same result was found for current users of pregnane derivatives as an HRT component (low quality evidence) compared to non-users.
- However, moderate quality evidence from a large cohort study (more than half million) showed an increased risk of VTE for users of medroxyprogesterone acetate as a component of combined HRT when compared with non-users.
- For non-testosterone derivatives as a component of combined HRT, moderate quality evidence from 1 large cohort study (more than half million) showed an increased risk of VTE in association to this type of HRT, whereas another study found no significant difference between this type of HRT and the no treatment group (very low quality evidence).

10.1.8 Evidence to recommendations

10.1.8.1 Relative value placed on the outcomes considered

The GDG considered VTE as one of the critical long term outcomes selected in this guideline development when considering the effect of HRT for women in menopause. VTE is associated with long term morbidity and via an increase in pulmonary embolus with increased mortality. This is a well-known complication that has been widely reported with use of sex steroid hormones (such as the combined oral contraceptive Pill) which impact on the liver synthesis of coagulation factors and thus increase the risk of clotting.

The GDG followed the principles set up at the Patient Experience Guideline (CG138) regarding the presentation of information to personalise risks and benefits as far as possible. For that purpose the use of absolute risk is preferred rather than relative risk. Information provision of all aspects of the benefit/risk ratio of HRT regarding short and long term is of paramount importance for women’s decision making regarding the choice of treatment for menopausal symptoms.

10.1.8.2 Consideration of clinical benefits and harms

Overall, evidence from both types of study designs, RCTs and observational studies, was largely consistent with regard to the increased risk of VTE associated with current oral HRT compared to non-users for women in menopause. The GDG concluded the following based on the main body of evidence:

Current oral HRT use was associated with a significant increased risk of VTE compared to non-users. Conversely, the risk of VTE was not found significantly different with transdermal HRT compared to non-users. This difference in the risk of VTE between oral and transdermal routes of HRT was supported by both RCTs and observational studies. In particular, subgroup analyses from observational studies showed that this trend also lasted for the duration of less than 2 years, less than 5 years, and more than 5 years of HRT use; and for women who started HRT use either before or after the age of 50 years. Furthermore, when
oral and transdermal HRT was compared head to head in two observational studies, both
studies found a significant increased risk of VTE among those on oral HRT compared to
those on transdermal HRT. Therefore, the group concluded that the information given to
women prior to HRT use should explain that the risk of VTE is increased with oral HRT use
whereas this is not the case for transdermal HRT. However, the group still wanted to draw
attention to women’s baseline risk of VTE and the recommendation explicitly stated that
transdermal HRT does not further increase the risk of VTE above the individual baseline
risk. Furthermore, due to the well-known VTE risk associated with obesity, the group decided
to give emphasis that the use of transdermal HRT is not contradicted for those with a high
BMI (over 30).

The included evidence showed that the increase in the VTE risk occurred rapidly after
starting HRT use and continued until treatment’s discontinuation. Evidence from randomised
participants showed a significantly increased risk of VTE within the first year of HRT use and
observational data on oral HRT also reported the same effect direction for up to two years’ of
HRT duration.

For women with previous episode of VTE, observational evidence showed no significant
difference between HRT users and non-users. However, the group considered that there
may be special considerations for this group of menopausal women before they start HRT
and they concluded that a referral to a haematologist should be offered.

Findings on different types of progesterone and progestogens in combined HRT were
inconclusive. Some observational studies showed an increased risk for some specific
preparation of progesterone or progestogen when combined with oestrogen, while other
studies found no significant difference. Therefore, the GDG decided not to differentiate the
direction of their decision making on HRT type.

Besides the general consistency in evidence, the GDG also noted the large sample sizes of
some included studies. For example 1 of those studies included more than half million
women. The GDG considered that although VTE was a significant side effect it was relatively
uncommon in those of menopausal age. It was found that 9 more per 1000 women (95%CI:
2 to 32 more) treated with HRT (oral or transdermal) may develop VTE within the first year of
use, and this absolute risk would increase to 10 more per 1000 women (95%CI: 5 to 13) for
the duration of 5 years.

10.1.8.3 Consideration of economic benefits and harms

VTE is expensive to manage and treat and has significant morbidity and mortality associated
with it. VTE is recognised as a potential adverse event arising from oral HRT treatment and
therefore it was important to consider it as part of an overall trade-off of risks and benefits of
therapy. This trade-off was done formally through an economic evaluation reported in
Appendix L although the analysis did not find that VTE outcomes were an important
determinant of cost effectiveness.

10.1.8.4 Quality of evidence

Moderate to very low quality evidence was presented in this review. For the included RCTs,
main reasons leading to downgrading of evidence was the small sample size, highly-selected
study population, open-label study design, high drop-out rates or disproportionate drop-out
rates between trial arms.

For the included comparative cohort studies, although some very large (half a million women)
and population based studies were included, findings from these studies were more likely
prone to risk of bias inherited from their study design. Major risk of biases shared by the
majority of included studies were: recall bias because of self-reported HRT use which could
result in misclassification of HRT use during follow-ups; different baseline characteristics in
HRT and no-treatment groups due to “self-selected” effect, generally HRT users were
healthier, younger, and had lower BMI; different confounders adjusted for in different studies, some studies could not adjust for some known risk factors such as family history of VTE and thrombophilia due to data availability in the analyses; different follow-ups for HRT and no-treatment groups or high drop-out rates but without reasons reported. Furthermore, imprecision of the risk estimates was another reason for which many outcomes were downgraded.

Evidence from high numbers of subgroup-analyses involved in some of the RCTs and observational studies should be interpreted with caution due to lack of precision in the estimation of effects.

10.1.8.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

The GDG discussed the importance of well-known risk factors for VTE such as age, genetic abnormalities, obesity, smoking or the presence of an inherited thrombophilia impacting on the clotting cascade with increase in coagulation (thrombophilias) and these should be taken into consideration when the prescription of HRT is considered. They also noted that some women with risk factors for VTE may be on anticoagulant therapy which means they may be considered for HRT with specialist advice.

The group also discussed the management of women who use HRT and are considered for elective surgery. Since transdermal HRT has little or no impact on coagulation and is not associated with an increased risk of VTE, there the group did not consider that there is a need of HRT discontinuation prior to elective surgery, especially when the surgery is minor and will not involve immobility. However, the group felt that this is a discussion that should take place between the woman, her surgeon and anaesthetist.

10.1.8.6 Key conclusions

The GDG concluded that:

- oral HRT (either oestrogen alone or oestrogen plus progesterone) increases the risk of VTE and this can occur immediately after starting HRT treatment
- there is no significantly increased risk of VTE in women using transdermal preparations compared to non-users
- there may be different impact on VTE risk of progesterone and the different progestogens on VTE risk when combined with oestrogen
- the background risk of VTE increases substantially with age and this should be taken into consideration when HRT use is considered for women in menopause
- the increased risk of VTE disappears after the HRT has been stopped

10.1.9 Recommendations

37. Explain to women that:
- the risk of venous thromboembolism (VTE) associated with HRT is greater for oral than transdermal preparations
- the risk associated with transdermal HRT given at standard therapeutic doses is no greater than baseline risk.

38. Consider transdermal rather than oral HRT for menopausal women who are at increased risk of VTE, including those with a BMI over 30.
39. Refer menopausal women at high risk of VTE (for example, those with a strong family history of VTE or a hereditary thrombophilia) to a haematologist for assessment before considering HRT.

### 10.1.10 Research Recommendations

<table>
<thead>
<tr>
<th>Research question</th>
<th>3. How does the preparation of HRT affect the risk of venous thromboembolism (VTE)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why this is needed</td>
<td>An increase in the risk of VTE (deep vein thrombosis [DVT] or pulmonary embolism [PE]) is a significant side effect of HRT, particularly as PEs can be fatal. This risk appears to be greater with oral than transdermal HRT. DVT risk increases with age and BMI, among other risk factors. The progestogen component of HRT may also influence the risk of a DVT, which may be greater with androgenic synthetic progestogens than natural progesterone (but findings from observational studies need confirmation). Most women in the UK take oral HRT comprising oestrogen combined with a synthetic progestogen, and the use of progesterone is less common. Randomised controlled trials are needed to compare oral with transdermal HRT, and HRT containing different types of progestogens. These trials should measure coagulation factors and the incidence of VTE in women at increased risk of VTE for whom transdermal oestrogen is indicated.</td>
</tr>
<tr>
<td>Importance to ‘patients’ or the population</td>
<td>Relevance to NICE guidance High importance. The recommendation in this guideline regarding VTE risk has been formulated using observational data from a relatively restricted population alone and this should be confirmed or amended as appropriate.</td>
</tr>
<tr>
<td>Relevance to the NHS</td>
<td>Reducing VTE risk would have a significant effect in regard to the safety of HRT overall and since there is also a suggestion that progesterone may have less impact on breast cancer risk, this could lead to a change in the type of HRT offered. This would make a difference to clinical practice as transdermal HRT would be recommended for women with risk factors e.g. raised BMI, and postmenopausal women initiating HRT could be offered oestrogen by the transdermal route in combination with progesterone as a first line treatment.</td>
</tr>
<tr>
<td>National priorities</td>
<td>N/A</td>
</tr>
<tr>
<td>Current evidence base</td>
<td>There are several studies quoted in the Guideline which include large epidemiological studies undertaken mainly in France. These are observational, thus the data are not conclusive.</td>
</tr>
<tr>
<td>Equality</td>
<td>N/A</td>
</tr>
<tr>
<td>Feasibility</td>
<td>It would be possible to conduct the study if the primary end point is an alteration in coagulation factors that provides an estimate of change in risk. A study of VTE event rate would be the “gold standard” but is likely to need inclusion of larger numbers to demonstrate a statistically significant difference. No other ethical or technical issues were identified.</td>
</tr>
<tr>
<td>Other comments</td>
<td>Traditional oral HRT might be contributed by the manufacturers of appropriate HRT preparations (e.g. oestradiol combined with norethisterone) and similarly transdermal HRT. Oestradiol and progesterone cannot be given by a single patch/gel but would require a combination of the patch/gel with either oral micronised progesterone or a vaginal pessary.</td>
</tr>
</tbody>
</table>
10.2 Cardiovascular disease

10.2.1 Review question
What are the effects of HRT administered for menopausal symptoms on the risk of development of cardiovascular disease (CVD) (including stroke) in women at different stages of the menopause?

10.2.2 Introduction to topic
CVD (including CHD and stroke) is the most common cause of death in women (1 in 2) worldwide. In certain parts of the world the rate of death due to CVD is known to be increasing.

There is a significant increase in the risk of developing CVD after the menopause, regardless of the age at which menopause occurs. Thus far, there has been controversy about the possible influence of HRT on CVD risk. Epidemiological data initially suggested that there might be a reduced risk of CHD with long term HRT usage. However, subsequent RCTs suggest that risk might be increased if hormone therapy is initiated at a later age. The aim of this review was to determine the precise CVD benefit/risk profile of hormonal products used during the menopause, thus empowering Health Care Providers and their patients to make fully informed therapeutic decisions.

10.2.3 Clinical introduction
The focus population of the review was women who have initiated treatment with HRT before the age of 65 years. Given that the risk of developing CVD may be different for women of different ages or at different menopausal stages, the Guideline Development Group decided at the protocol stage to produce subgroup analyses on the following: age distribution, user categories (i.e. ever, past and current users), durations of HRT use, timing of HRT initiation relative to the onset of menopause, time since stopping HRT, different treatment administration routes and different preparations of HRT.

RCTs and comparative prospective cohort studies were selected for inclusion in this review. For comparative cohort studies, only those that have adjusted their analyses for the most common confounders such as age, hypertension, BMI were included.

For full details see review protocol in Appendix D.

10.2.4 Description of included studies
Five RCTs comparing some form of HRT with control or placebo group were included (Table 18):
- one trial (Shierbeck 2012) conducted in Denmark examined the effect of HRT on coronary disease and stroke in comparison with control group;
- one trial (the Women’s Health Initiative [WHI], which included 9 intervention or post-intervention reports conducted in the USA: Anderson 2004; Hendrix 2006; Lacroix 2009; Manson 2003; Manson 2013; Prentice 2009; Rossouw 2007; Toh 2010; Wassertheil-Smoller 2003). It examined the effect of oestrogen alone, and oestrogen plus progesterogen on CHD and stroke in comparison with placebo groups;
- one study (Cherry 2014) conducted in the UK between 2000 and 2002 examined the effect of oestrogen on ischemic heart disease death among women with an intact uterus in comparison with placebo. The included report focused on the IHD death outcome during its post-intervention phase where HRT use or not could not be ascertained after the active intervention was finished;
- two trials conducted in the USA assessed the effect of HRT on blood pressure change in comparison with placebo (Brownley 2004; The Writing Group for the PEPI Trial, 1995)

### Table 18: Main characteristics of the RCTs included in the review

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Age in years (mean or range)</th>
<th>Sample size: HRT/control or placebo</th>
<th>HRT type</th>
<th>Outcome</th>
<th>Duration of intervention in years (mean, median), (post-intervention follow-up if existing)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson 2004 (also reported in Hendrix 2006; Lacroix 2011)</td>
<td>USA</td>
<td>50-59</td>
<td>1637/1673</td>
<td>Oestrogen</td>
<td>CHD; Stroke</td>
<td>Mean 6.8 years</td>
</tr>
<tr>
<td>Manson 2002 (also reported in Wassertheil-Smoller 2003)</td>
<td>USA</td>
<td>50-59</td>
<td>1223/1232</td>
<td>Oestrogen</td>
<td>CHD; Stroke</td>
<td>Median 5.9 years post-intervention;</td>
</tr>
<tr>
<td>Manson 2013</td>
<td>USA</td>
<td>50-59</td>
<td>N/R; N/R</td>
<td>Oestrogen plus progestogen and Oestrogen alone</td>
<td>CHD; Stroke; MI</td>
<td>Median 8.2 years post-intervention for CEE trials; Median 6.6 years post-intervention for CEE plus MPA trials;</td>
</tr>
<tr>
<td>Prentice 2009</td>
<td>USA</td>
<td>HRT initiated within 2. 2 to 4 years since menopause</td>
<td>N/R</td>
<td>Oestrogen plus progestogen and Oestrogen alone</td>
<td>CHD; Stroke</td>
<td>Data from two WHI clinical trials and observational trials are combined; CEE plus MPA trial: mean 5.2 years; CEE trial: mean 6.8 years; The observational cohorts of WHI: CEE cohort: mean 7.1 years; CEE plus MPA cohort: 5.5 years, respectively.</td>
</tr>
<tr>
<td>Rossouw 2007</td>
<td>USA</td>
<td>50-59 years; HRT initiated within 10 years since menopause</td>
<td>N/R</td>
<td>HRT (the two WHI clinical trials combined)</td>
<td>CHD; Stroke</td>
<td>CEE plus MPA trial: mean 5.2 years; CEE trial: mean 6.8 years;</td>
</tr>
<tr>
<td>Cherry 2014</td>
<td>UK</td>
<td>50-59 years</td>
<td>167/134 (participants were women who have survived an MI)</td>
<td>Oestrogen (oestradiol valerate)</td>
<td>IHD death</td>
<td>Mean 12.6 years including both intervention (2 years) and post-intervention phases.</td>
</tr>
<tr>
<td>Toh 2010</td>
<td>USA</td>
<td>50-59 years, less than and more than 2 years duration</td>
<td>2839/2683</td>
<td>Oestrogen plus Progestogen</td>
<td>CHD; Stroke</td>
<td>Mean 5.2 years</td>
</tr>
<tr>
<td>Brownley 2004</td>
<td>USA</td>
<td>50.6 ± 0.9</td>
<td>19/23</td>
<td>HRT</td>
<td>Blood pressure change</td>
<td>6 months</td>
</tr>
<tr>
<td>The Writing Group for the PEPI Trial, 1995</td>
<td>USA</td>
<td>56.1</td>
<td>701/104 174</td>
<td>CEE; CEE, 0.625 mg/d, plus MPA, 10 mg/d for first 12 days: N = 174 CEE, 0.625 mg/d, plus MPA, 2.5 mg/d: N = 174; CEE, 0.625 mg/d, plus MP, 200 mg/d for first 12 days: N</td>
<td>Blood pressure change</td>
<td>3 years</td>
</tr>
</tbody>
</table>

**National Collaborating Centre for Women's and Children's Health**

132
Eighteen cohort studies comparing HRT use with no use were included (Alexander 2001; Corrao 2007; Graff-Iversen 2004; Gast 2011; Ettinger 1996; Folsom. 1995; Hedblad 2002; Hernandez 1990; Li 2006; Lokkegaard 2008; Sourander 1998; Lafferty 1994; Pentti 2006; Shlipak 2001; Stram 2011; Su 2012; The Nurses’ Health Study (The NHS) with 5 publications; Grodstein 1996; Grodstein 2000; Grodstein 2006; Grodstein 2008; Stamper 1985, Weiner 2008).

The majority of the cohort studies were conducted in the USA, some in Europe (UK, Denmark, Finland, the Netherlands, Italy), and one in Chile and Turkey. Sample sizes varied and ranged from 157 (Lafferty 1994) to 698,098 participants (Lokkegaard 2008).

A summary of the cohort studies that were included in this review are presented in Table 19.

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Age in years (mean or range)</th>
<th>Sample size</th>
<th>HRT type</th>
<th>Study follow-up</th>
<th>Confounders in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexander 2001</td>
<td></td>
<td>59 (52.66)</td>
<td>1857</td>
<td>HRT</td>
<td>2 years</td>
<td>Age, previous angina, congestive heart failure, current smoker, hypertension, prior MI, PVD, prior stroke or TIA, race, weight, and previous randomised treatment</td>
</tr>
<tr>
<td>Corrao 2007</td>
<td></td>
<td>54.7</td>
<td>88,050</td>
<td>HRT</td>
<td>3 years</td>
<td>Exposures to cardiac drugs, antihypertensives, lipid modifying agents, drugs used in diabetes, raloxifene, and other sex hormones during follow-up</td>
</tr>
<tr>
<td>Graff-Iversen 2004</td>
<td></td>
<td>35-62</td>
<td>14,324</td>
<td>HRT</td>
<td>14 years</td>
<td>Age and CVD health</td>
</tr>
<tr>
<td>Gast 2011</td>
<td></td>
<td>46-64</td>
<td>8,865</td>
<td>HRT</td>
<td>10 years</td>
<td>Age, education level, smoking, physical activity, hypertension, hypercholesterolemia, menopausal status, and oral contraceptive use</td>
</tr>
<tr>
<td>Ettinger 1996</td>
<td>Women within 3 years of menopause</td>
<td></td>
<td>454</td>
<td>HRT</td>
<td>26.8 years</td>
<td>Age, BMI, smoking, alcohol consumption, hypertension, abnormal ECG, and total serum cholesterol level above 260 mg/dL</td>
</tr>
<tr>
<td>Folsom 1995</td>
<td></td>
<td>55-59</td>
<td>41,837</td>
<td>HRT</td>
<td>6 years</td>
<td>Age, marital status, physical activity level, alcohol use, smoking, BMI, waist/hip ratio, hypertension, and diabetes</td>
</tr>
<tr>
<td>Grodstein 1996</td>
<td></td>
<td>58.5</td>
<td>59,337</td>
<td>HRT; oestrogen;</td>
<td>16 years</td>
<td>Age, age at menopause, BMI, smoking, hypertension, diabetes, elevated cholesterol levels, myocardial infarction in a parent before the age of 60 years, prior use of oral contraceptives, type of menopause, and two-year interval</td>
</tr>
<tr>
<td>Grodstein 2000</td>
<td></td>
<td>30-55</td>
<td>70,533</td>
<td>HRT; Oestrogen;</td>
<td>20 years</td>
<td>Age, BMI, history of diabetes, hypertension, high cholesterol level, age at menopause, smoking, and parental history of premature heart disease</td>
</tr>
<tr>
<td>Grodstein 2006 (the NHS)</td>
<td></td>
<td>30-55</td>
<td>121,700</td>
<td>HRT; Oestrogen;</td>
<td>24 years</td>
<td>Age, BMI, smoking, history of hypertension, elevated cholesterol, parental MI before age 60. For certain analyses, husband’s education was also adjusted for as an additional measure of socioeconomic status.</td>
</tr>
<tr>
<td>Grodstein 2008</td>
<td></td>
<td>30-55</td>
<td>121,700</td>
<td>HRT; Oestrogen;</td>
<td>28 years</td>
<td>Age, BMI, height, smoking, history of hypertension, diabetes, and elevated cholesterol level, husband’s education, and parental MI before the age of 60 years.</td>
</tr>
<tr>
<td>Hedblad 2002</td>
<td></td>
<td>53.8</td>
<td>5,721</td>
<td>HRT</td>
<td>9.2 years</td>
<td>Age, BMI, hypertension, diabetes,</td>
</tr>
<tr>
<td>Study</td>
<td>Age in years (mean or range)</td>
<td>Sample size</td>
<td>HRT type</td>
<td>Study follow-up</td>
<td>Confounders in analysis</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------</td>
<td>-------------</td>
<td>----------</td>
<td>-----------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Hernandez 1990</td>
<td>50-64</td>
<td>310,000</td>
<td>HRT</td>
<td>6 years</td>
<td>Age in 5-yr intervals and for period in 2-year intervals</td>
<td></td>
</tr>
<tr>
<td>Li 2006</td>
<td>56</td>
<td>16,906</td>
<td>HRT</td>
<td>10.5 years</td>
<td>Age, smoking, alcohol consumption, BP, BMI, diabetes, use of BP lowering agents, lipid-lowering agents or aspirin</td>
<td></td>
</tr>
<tr>
<td>Lokkegaard 2008</td>
<td>51-64</td>
<td>698,098</td>
<td>HRT</td>
<td>6 years</td>
<td>Age, calendar year, education, employment status, habitation, medication for hypertension, heart conditions, hyperlipidaemia, or diabetes;</td>
<td></td>
</tr>
<tr>
<td>Sourander 1998</td>
<td>60</td>
<td>7,944</td>
<td>HRT</td>
<td>7 years</td>
<td>Social class, smoking, age, BMI, diabetes, hypertension, CVA, and cardiac failure</td>
<td></td>
</tr>
<tr>
<td>Lafferty 1994</td>
<td>53</td>
<td>157</td>
<td>HRT</td>
<td>14 years</td>
<td>Age only</td>
<td></td>
</tr>
<tr>
<td>Pentti 2006</td>
<td>57.3</td>
<td>11,667</td>
<td>HRT</td>
<td>7 years</td>
<td>Age, parity, BMI, hysterectomy, bilateral oophorectomy, number of chronic health disorders and time since menopause</td>
<td></td>
</tr>
<tr>
<td>Shlipak 2001</td>
<td>55-64</td>
<td>114,724</td>
<td>HRT</td>
<td>2 years</td>
<td>Age, race, diabetes, hypertension, smoking, hypercholesterolemia, prior MI, prior stroke, prior angina, prior heart failure, presence of chest pain, time to presentation to hospital, BP, heart rate, admission diagnosis etc.</td>
<td></td>
</tr>
<tr>
<td>Stram 2011</td>
<td>36-64</td>
<td>71,237</td>
<td>HRT</td>
<td>5 to 7 years</td>
<td>BMI, smoking status, alcohol consumption, physical activity, total caloric intake, and cholesterol during the year before baseline, Self-reported history of diabetes, high blood pressure, MI or heart disease, cancer and stroke</td>
<td></td>
</tr>
<tr>
<td>Su 2012</td>
<td>≤ 55</td>
<td>16,045</td>
<td>HRT</td>
<td>7 years</td>
<td>Age, statin use, aspirin use, hypercholesterolemia, diabetes medication use and hypertension</td>
<td></td>
</tr>
<tr>
<td>Stampfer 1985</td>
<td>30-55</td>
<td>121,964</td>
<td>HRT</td>
<td>4 years</td>
<td>Age and other potential risk factors</td>
<td></td>
</tr>
<tr>
<td>Weiner 2008</td>
<td>52.3</td>
<td>26,536</td>
<td>HRT</td>
<td>9 years</td>
<td>BP, BMI, and smoking</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** BMI - body mass index; BP - blood pressure; CVD - cardiovascular disease; HRT – hormone replacement therapy; MI - myocardial infarction; TIA - transient ischemic attack

### 10.2.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F; study evidence tables in Appendix H; forest plots in Appendix J; and exclusion list in Appendix G.

### 10.2.6 Economic evidence

No health economic search was undertaken for this guideline as the decision was made to prioritise short term treatment. 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. There was evidence that HRT increases the risk of stroke when administered orally, however the absolute risk was very small and therefore the clinical evidence from this review was not used to inform the model on short term treatment.
10.2.7 Evidence statements

Evidence statements for RCTs

Low to very low quality evidence from 1 RCT of over 1000 women (mixed population of women with and without a uterus) aged 45-58 years who were followed up for 10 years found that:

- the risk of CHD is significantly lower for those women in the HRT group compared with placebo. This beneficial effect of HRT followed the same direction in the subgroup analysis for the age groups of 45-49 and 50-58 years old and was this effect was preserved for 6 years after HRT termination.
- no difference was found for the outcome of stroke whereas some indication was found for reduction of systolic blood pressure with the use of HRT in a different RCT (of 42 women).

Low to very low quality evidence from over 5000 women in a RCT with post-intervention follow-up found no significant difference between oestrogen plus progestogen and placebo users in the risk of CHD or stroke among women aged between 50 and 59 years.

Low quality evidence from over 3000 women in the same RCT with post-intervention follow-up found no significant difference between oestrogen and placebo users in the risk of CHD or stroke among women aged between 50 and 59 years.

Reanalyses of the previously mentioned RCT (with concluded the following findings: (of no difference in the CHD or stroke risk between oestrogen plus progestogen and no HRT users (or placebo) specifically for:

- women treated with oestrogen plus progestogen for more or less than 2 years had no significant difference in the CHD or stroke risk when compared to no HRT users (very low quality evidence)
- women who initiated oestrogen plus progestogen within 2, 2-4, 5 years or 10 years of menopause (with and without prior HRT use) had no significant difference in the CHD or stroke risk when compared to no HRT users (low to very low quality evidence). This finding was also applied to women who initiated HRT within 10 years since menopause and with a duration of less or more than 2 years (very low quality evidence)
- women aged 50-59 years at baseline, there was no significant difference between those who had oestrogen plus progestogen and no HRT in the risk of CHD, stroke or total MI after a median 8.2 years following termination of the therapy (median cumulative follow-up 13.2 years) (low quality of evidence)
- women who initiated oestrogen within 2, 2-4, 5 or 10 years of menopause (with and without prior HRT use) had no difference in the CHD or stroke risk compared to non-users (low quality of evidence)
- women aged 50-59 years at baseline there was no significant difference in the risk of stroke between those who had oestrogen compared with no HRT after a median of 5.9 or 6.6 years since the therapy’s termination (low quality of evidence) women aged 50-59 years at baseline had a significantly reduced risk of CHD or total MI after a median of 5.9 years from termination of the oestrogen therapy compared with no HRT. The same significantly reduced risk was found in another post-intervention reanalysis of the trial for after a median 6.6 years since termination of the therapy (low quality of evidence).

When the two clinical trials (oestrogen alone, oestrogen plus progestogen) of the previous RCT were combined in another reanalysis, no significant difference was found in the risk of CHD and stroke for women aged 50-59 years at baseline compared to non users. However, when data in this reanalysis were analysed according to time of initiating HRT since menopause started, a significantly increased risk of stroke was found among women who initiated HRT use within 10 years since menopause but no significant difference was found for the risk of CHD compared to non users. The quality of the evidence was very low.
A post-intervention analysis of one RCT which examined the effect of oestrogen on women with an intact uterus who have survived an MI, found no significant difference in the risk of IHD death after a mean 10 years follow-up between those who had oestrogen and no HRT (very low quality evidence).

For the outcomes of blood systolic and diastolic blood pressure, moderate to low quality evidence from one RCT found no significant difference in either systolic or diastolic pressure among HRT users of different preparations (oestrogen, oestrogen plus MPA cyclic, oestrogen plus MPA daily) compared with the placebo group. However, another RCT showed that there was a significant reduction in the mean systolic and diastolic pressure among HRT users of less than 5 years duration compared with the placebo group.

Evidence statements for comparative cohort studies

Risk of CHD in relation to HRT use according to user category

Very low quality evidence of a meta-analysis of 4 cohort studies with more than 70,000 participants showed a significant reduction in the risk of CHD between current HRT and non-users.

However, a subgroup analysis of two cohorts in women younger than 55 years old found no difference in the risk of CHD among current and non HRT users (very low quality evidence).

Further analysis showed that:

- there was a significant reduction in the risk of CHD among those who have used HRT for durations of more than 2 or 5 years compared to non-users (very low quality evidence from two cohorts)
- however, moderate quality evidence from another cohort study found no difference in the risk of CHD (defined as cardiac events-composite of death/MI/unstable angina [UA]) or MI in current and prior users with more than 2 years of duration
- no difference was found in the risk of CHD among ever HRT users with or without the presence of flushing symptoms. The evidence was low to very low quality and came from one cohort study
- a significantly higher risk of CHD in current HRT users with pre-existing heart disease was found compared to non-users. The evidence was low quality and from two cohorts

For the outcome of IHD different cohorts came to the following conclusions:

- one cohort study found no significant difference in the risk of IHD among users of any route of HRT with a duration of 7 to 12 months compared with HRT users of less than 6 months’ duration. The same direction of effect was found in users of any route of HRT with a duration of 1 to 2 years and 2 to 3 years. The quality of this evidence was low. Low to very low quality evidence that had the different routes of administration (oral, transdermal) supported the same conclusion

For the outcome of death from IHD, CVD or CHD, the following conclusions were drawn:

- very low quality evidence from single cohort studies found no difference in the risk of IHD death in past HRT users aged 36-59 years compared with non-users. The same was found in past HRT users aged 60-64 years and for HRT users who initiated the treatment at the age of 45 to 54 years or 55 to 64 years
- timing of initiation of HRT since menopause was not found to impact on the previous finding that there is no difference in the risk of ischemic heart disease death in women who initiated HRT use within 5 or 10 years since menopause compared with non-users
- meta-analysis of 4 cohort studies showed a significantly lower risk of CVD death in current HRT users compared with non-users. The quality of evidence was low
• meta-analysis of 4 cohort studies showed a significantly lower risk of CHD death in current
HRT users compared with non-users. The quality of evidence was very low
• very low quality evidence from two cohort studies found no difference in the risk of CHD
death in current HRT users of more than 5 years' duration compared with non-users

For the outcome of total stroke (generally including fatal and non-fatal, ischemic and
haemorrhagic stroke in studies):

Low quality evidence from different cohorts comprising more than 50,000 participants
showed the following results:
• a significantly increased risk of total stroke in current HRT users compared with non-users
• no difference in the risk of total stroke in current HRT users with a duration of more than 2
(2 cohorts) or 5 years (2 cohorts) compared with nonusers. The quality of evidence was very low
• no difference in the risk of stroke among users of any route of HRT and with a duration of
7 to 12 months compared with HRT users of less than 6 months duration. The same was
found in users of any route of HRT with a duration of 1 to 2 years. However, a significantly
reduced risk of stroke was found in users of any route of HRT with a duration of 2 to 3
years, and of more than 3 years when compared with HRT users of less than 6 months
duration. The quality of evidence was low
• one cohort study found a significantly reduced risk of stroke among users of transdermal
HRT and with a duration of 7 to 12 months, 2 to 3 years, and more than 3 years when
compared with HRT users of less than 6 months duration. However, among users of
transdermal HRT with a duration of 1 to 2 years, no difference was found in the risk of
stroke when compared with oral HRT users of less than 6 months duration. The quality of
evidence was low
• one cohort study found no difference in the risk of stroke among users of oral HRT with a
duration of 7 to 12 months compared with HRT users of less than 6 months duration. The same
was found in users of oral HRT with a duration of 1 to 2 years, 2 to 3 years, and
more than 3 years. The quality of evidence was low to very low

For systolic blood pressure, very low quality evidence from one cohort showed that there is
no significant difference in mean values between HRT users and non-users at 6-month
follow-up, whereas the same evidence showed that there was a significant decrease in
diastolic blood pressure for HRT users.

Low to very low quality evidence from a large prospective cohort study is summarized below:
• significantly reduced risk for total CHD among current HRT users compared with non-
users at 4-, 10-, 16-, and 20- years of follow-up
• no significant difference in the risk for CHD between current HRT users and non-users
among women aged less than 50 years. However, among women aged between 50-59
years, a significantly reduced risk was found among current HRT users compared with
non-users
• a significantly reduced risk for CHD among current HRT users of less than 1 year duration
compared with nonusers. This reduced risk was also shown for duration 1 to 2 years, 2 to
4.9 years, 5 to 9.9 years, and more than 10 years
• a significantly reduced risk for total CHD among oestrogen users compared with non-
users at 24-year follow-up. Also at 24-year follow-up when women with and without pre-
existing heart disease were included in the analysis
• a significantly reduced risk of total CHD among oestrogen plus progestogen users
compared with nonusers at 16 and 24-year follow-up. This remained significant when
women with and without pre-existing heart disease were included in the analysis at 24-
year follow-up
Menopause
Long-term benefits and risks of HRT

- no difference in the risk for total CHD between past HRT users and nonusers at 4-, 10-, and 6-year follow-up
- a significantly reduced risk for CHD among past users of HRT compared with non-users at 20-year follow-up

Subgroup analysis based on the age distribution found that:
- among women aged between 40 and 44 years, there was no significant difference in the risk for total CHD in either ever or current users compared with non-users. However, among women aged between 45 and 49 years, a significantly reduced risk for total CHD was found in both ever and current users compared with non-users
- among women aged between 50 and 55 years, the same study found no significant difference between ever and non-users, while a significantly reduced risk was found among current users compared with non-users
- among women aged 56 and 59 years, the same study showed no significant difference in the CHD risk between ever users and non-users

Low to very low quality evidence from reanalyses of cohort studies found:
- a significantly reduced risk of CHD among women who initiated oestrogen or oestrogen plus progestogen use within 4 years of menopause compared with non-users. However, a non-significant difference was found between those who initiated oestrogen at least 10 years after menopause compared with non-users
- a significantly reduced risk of non-fatal MI among current HRT users compared with non-users at 4-year follow-up but not between past and non-users
- no difference in the risk of CVD between current or past HRT users and non-users at 10-year follow-up
- a significantly reduced risk for CVD death in current but not in past HRT users compared with non-users at 16-year follow-up
- no difference in the risk of total stroke between current HRT users and non-users at 10-, 16-, and 20-year follow-up. The conclusion was reached when the total stroke was broken down into ischaemic and subarachnoid stroke at 10-year follow-up
- at 16-year follow-up, a significantly increased risk for ischemic stroke was found among current users compared with non-users, while no difference was found in subarachnoid stroke among current HRT users and non-users
- among current HRT users of less than 1 year duration, no significant difference in the risk of stroke was found between current HRT users and non-users. This non-difference was also found for the durations of 1 to 2 years, 2 to 4.9 years, 5 to 9.9 years, and more than 10 years
- no difference was found in the risk of total stroke between current oestrogen or oestrogen plus progestogen users and non-users at 16-year follow-up
- a significantly increased risk of stroke (including ischemic) among current oestrogen or oestrogen plus progestogen users compared with non-users at 28-year follow-up. The risk of haemorrhagic stroke was found significantly increased for current oestrogen users when compared to non-users but not for current oestrogen plus progestogen users and non-users at 28-year follow-up
- a significantly increased risk of stroke among women who initiated oestrogen use within 4 years, and at least 10 years after menopause but not for those who initiated oestrogen plus progestogen compared with non-users. The quality of evidence was very low
- no significant difference between women who initiated HRT within 4 years, or at least 10 years after menopause, compared with non-users. A significantly increased risk for stroke among women who initiated oestrogen use at age 50 to 59 years. However, among women of the same age who initiated oestrogen plus progestogen, no significant difference was found
no difference in the risk of stroke or fatal stroke death between current or past HRT users (either oestrogen alone or oestrogen plus progestogen) and non-users at 16-year or 28 year follow-up

10.2.8 Evidence to recommendations

10.2.8.1 Relative value placed on the outcomes considered

The GDG considered different types of CVD such as stroke and MI, cardiac event composite scores, change in blood pressure and mortality from CVD as the most important outcomes for this review question. The GDG followed the principles outlined in the Patient Experience Guideline (CG138) regarding the presentation of information to personalise risks and benefits as far as possible. For that purpose, the use of absolute risk is preferred rather than relative risk. Information provision of all aspects of the benefit/risk ratio of HRT regarding short and long term consequences of treatment is of paramount importance for women’s decision making regarding the choice of treatment for menopausal symptoms (linked to other long term symptom reviews).

10.2.8.2 Consideration of clinical benefits and harms

Randomised evidence from several thousand women aged between 45 and 58 years old consistently showed that the risk of stroke and MI is not significantly different between menopausal women who received HRT (either as oestrogen alone or as a combination of oestrogen plus progestogen) and those who received no treatment.

Subgroup analyses of RCT data also showed an absence of harm for those women being treated with either oestrogen alone or oestrogen plus progestogen and this was preserved independently of the timing of initiation of HRT (within 2-, 4-, 5- or 10- years since menopause) and duration of HRT. This result also remained 6 or 8 years after termination of HRT.

Evidence from observational studies revealed similar conclusions to those drawn from RCTs, although more information was provided for specific subgroups (for example women with pre-existing heart disease), different routes of HRT administration and different HRT durations.

The GDG placed more emphasis on the following results from the observational studies when they were drafting the recommendations:

- the risk of CHD was significantly lower for women using HRT compared with no treatment across different follow-up periods (4-, 10-, 16- and 20- years) and different HRT durations (1-, 2-, 5- or 10- years) although the risk seemed to significantly increase in current users with pre-existing heart disease
- conflicting results were found as to whether the risk of CVD or CHD is reduced or is similar in current HRT users compared to non-users
- some observational data found that the risk of stroke may be higher for women under the age of 55 years who are on HRT compared to non-users, whereas other evidence found no difference in the outcome of stroke among users of any route of HRT with different duration of use and long term follow-up (16, 20 years) when compared to non-users.
- weak data suggesting transdermal HRT administration may be associated with a lower risk of stroke than oral

The GDG discussed the role of age in development of heart disease; CHD risk rises for everyone as they age, but for women specifically, cardiovascular symptoms can become more evident after the onset of menopause. Although menopause does not cause CVD, there may be associated risk factors (such as smoking, poor diet, lack of exercise) that increase the risk of CVD around the time of menopause. The GDG considered in details the
synthesis of evidence and they concluded that there is no clear evidence of harm in terms of CHD or stroke in menopausal women under the age of 65 years who are taking HRT and when HRT is terminated. Therefore, there is enough evidence to support health professionals in advising women of the low or no risk in CVD outcomes associated with the use of HRT. In addition, although there were limited data indicating that there may be a significant increase in CHD found in current HRT users with pre-existing conditions compared to non-users, the group did not feel that this evidence was compelling enough to draft a negative recommendation for information giving.

Based on UK data, the baseline risk of CVD and stroke is low: 26.3 per 1000 and 11.3 per 1000 (Weiner 2008) respectively, over 7.5 years, which increases with age, but is not significantly increased by the use of HRT.

10.2.8.3 Consideration of economic benefits and harms

The evidence shows that HRT increases the risk of stroke of women who are in the menopause. However, the absolute risk is very small and therefore the economic benefits and harms are limited. There is a suggestion that transdermal preparations have less impact on the risk of stroke than oral preparations.

10.2.8.4 Quality of evidence

The majority of RCT evidence was low to very low quality, largely due to high risk of bias (mainly due to unblinding of study design) and the lack of confidence in the direction of effect size (imprecision). The WHI data, which contributed substantially to the RCT evidence base, had some design limitations: the study included a group of healthy menopausal women with a high baseline BMI (35-40% of this group had BMI 30 or over) and was terminated earlier than expected due to high prevalence of side effects. In addition, a proportion of the women included in the trial had initiated treatment outside of the study’s protocol (for example 9.1% in the placebo arm were using HRT) and 36% of them had previous HRT experience. Thus the greatest concern in using the WHI study was the external validity of the estimates given by the characterisation of the present study population. Furthermore, the information from the post-intervention period is unblinded. Several post-hoc analyses have been included for the presentation of the relevant evidence and the results of these analyses should be interpreted with caution due to lack of statistical power in these type of analyses. However, the sample size of this trial was sufficiently large to allow clinically relevant conclusions.

The majority of observational evidence (cohort studies) were of very low quality. The main methodological limitations of these studies were the difference in baseline characteristics between the HRT and no treatment arms, the highly selective approach of the included population (for example the Nurse’s Health Study included only nurses (potentially a healthy cohort) and the serious heterogeneity and imprecision observed in some of the results. Given that these data were observational and the role of confounding factors is important on the estimates of effects, evidence was downgraded if the results were not adjusted for the most relevant confounders (such as age and HRT duration).

10.2.8.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

This review question looked at the impact of HRT use, duration, timing since stopping and age on the CVD risk but did not consider any potential differences in outcome related to the different formulations or the type and dosage of HRT in the preparations, although the clinical experience of the GDG suggests that there may be differential effects and further research in this area is needed.
Although the GDG concluded that menopausal women should be informed that the risk of CHD associated with HRT use is low or minimal, they highlighted the need for all women around the age of menopause to have their personal cardiovascular risk reviewed on an ongoing basis as part of in line with NICE guidance 181 on lipid modification.

### 10.2.8.6 Key conclusions

The GDG concluded that:

- there is no convincing evidence that the administration of HRT increases the risk of CVD in women under 65 years of age. This is evidence for both oestrogen and oestrogen plus progestogen preparations and is not influenced by route of administration.
- there is evidence to show that HRT increases the risk of stroke in women in the menopause. There is a suggestion that transdermal preparations have less impact on the risk of stroke than oral preparations.
- there is no evidence of increased risk of haemorrhagic stroke with HRT administration.

### 10.2.9 Recommendations

40. **Ensure that menopausal women and healthcare professionals involved in their care understand that HRT:**
   - does not increase cardiovascular disease risk when started in women aged under 60 years
   - does not affect the risk of dying from cardiovascular disease.

41. **Be aware that cardiovascular risk factors (for example hypertension) do not automatically preclude a woman from taking HRT but should be taken into account.**

42. **Using tables 1 and 2, explain to women that:**
   - the baseline risk of coronary heart disease and stroke for women around menopausal age varies from one woman to another according to the presence of cardiovascular risk factors.
   - HRT with oestrogen alone is associated with no, or reduced, risk of coronary heart disease.
   - HRT with oestrogen and progestogen is associated with little or no increase in the risk of coronary heart disease.

43. **Explain to women that taking oral (but not transdermal) oestrogen is associated with a small increase in the risk of stroke. Also explain that the baseline risk of stroke in women aged under 60 years is very low (see table 2).**

#### Table 1: Absolute rates of CHD for different types of HRT compared with no HRT (or placebo), different duration of HRT use and time since stopping HRT for menopausal women

<table>
<thead>
<tr>
<th></th>
<th>Difference in coronary heart disease incidence per 1000 menopausal women over 7.5 years (baseline risk in the UK population over 7.5 years: 263 women per 1000 [Weiner et al. 2008])</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Past users</td>
</tr>
<tr>
<td><strong>Women on oestrogen alone</strong></td>
<td></td>
</tr>
<tr>
<td>RCT estimate’</td>
<td>–</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
</tr>
</tbody>
</table>
### Table 2: Absolute rates of stroke for different types of HRT compared with no HRT (or placebo), different duration of HRT use and time since stopping HRT for menopausal women

<table>
<thead>
<tr>
<th>Past users</th>
<th>Current users</th>
<th>Treatment duration &lt;5 years</th>
<th>Treatment duration 5–10 years</th>
<th>&gt;5 years since stopping treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women on oestrogen alone</td>
<td>RCT estimate¹</td>
<td>–</td>
<td>1 more (from 5 fewer to 14 more)³</td>
<td>–</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
<td>3 more (from 1 fewer to 8 more)²</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Women on oestrogen plus progestogen</td>
<td>RCT estimate¹</td>
<td>–</td>
<td>5 more (from 3 fewer to 20 more)³</td>
<td>–</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>–</td>
<td>4 more (from 1 more to 7 more)²</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Women on any HRT</td>
<td>RCT estimate</td>
<td>–</td>
<td>3 fewer (from 7 fewer to 8 more)³</td>
<td>–</td>
</tr>
<tr>
<td>Observational estimate</td>
<td>0 fewer (from 2 fewer to 2 more)²</td>
<td>3 more (from 2 more to 5 more)²</td>
<td>–</td>
<td>1 more (from 2 fewer to 4 more)²</td>
</tr>
</tbody>
</table>

HRT, hormone replacement therapy; RCT, randomised controlled trial

1. For women aged 50–59 years
2. Anderson 2004 (the WHI)
3. Lacroix 2011 (the WHI reanalysis)
4. Grodstein 1996 (the NHS)
5. Manson 2003 (the WHI)
6. Manson 2013 (the WHI reanalysis)
7. Schierbeck 2012
8. Grodstein 2000 (the NHS)
9. Hedblad 2002; Lokkegaard 2008; Stram 2011; Grodstein 2000 (the NHS)
10. Folsom 1995; Grodstein 2000 (the NHS)
11. For women aged 50–59 years
12. Anderson 2004 (the WHI)
13. Lacroix 2011 (the WHI reanalysis)
14. Grodstein 1996 (the NHS)
15. Manson 2003 (the WHI)
16. Manson 2013 (the WHI reanalysis)
17. Grodstein 2000 (the NHS)
18. Schierbeck 2012
19. Grodstein 2000 (the NHS); Li 2006; Sourander 1998
20. Grodstein 2000 (the NHS)
21. Folsom 1995; Grodstein 2000 (the NHS)
22. Anderson 2004 (the WHI)
10.3 Development of Type 2 Diabetes

10.3.1 Review question

What are the effects of HRT administered for menopausal symptoms on the risk of developing type 2 diabetes (T2DM)?

10.3.2 Introduction to topic

More than 3% of the UK population have T2DM, with rates rising to 5-7% in areas where larger proportions of the population are of South Asian or African/Caribbean origin. In addition to genetic factors, including family history, increasing age is an important risk factor for T2DM as is abnormal glucose tolerance (impaired fasting glycaemia, IFG). Rates of IFG increase from 15.3% for women aged 40-49 years to 28.1% aged 60-69 years, as the incidence of T2DM increases from middle-age onwards.

Insulin resistance and pancreatic beta cell depletion are common features of T2DM. While androgens reduce peripheral insulin sensitivity, oestrogens antagonise this effect. At the menopause, reduced levels of oestrogen with relatively increased androgenic activity may result in impaired glucose tolerance and central obesity, possibly explaining the increased T2DM risk (Collins 2007).

Dyslipidaemia is an important component of T2DM and there is a 3 to 5 times greater risk of death from IHD among diabetic women. Changes in serum lipids and lipoprotein profiles are seen at the menopause, with increases in serum triglycerides and low density lipoproteins but decreasing high density lipoproteins especially HDL2 subfractions (Collins 2007)

Although women with type 1 diabetes have better lipid profiles than women with T2DM, by the age of the menopausal their incidence of IHD is 9 times higher than that of non-diabetic women, probably due to endothelial dysfunction and microvascular changes.

Thus, maintaining physiological oestrogen levels in postmenopausal women could be hypothesised to decrease the incidence of abnormal glucose tolerance, T2DM and associated dyslipidaemia with the potential to improve cardio-metabolic risk.

10.3.3 Clinical introduction

The aim of this review was to assess the effect of HRT use on the risk of developing T2DM in menopausal women. Subgroup analyses of the age distribution of the participants or on the stage of menopause (peri or post menopause) were presented if data were available.

The risk of developing T2DM was examined in terms of different HRT types, current or past HRT use, duration and timing since stopping if data were available.

Given the interventional nature of this review question, we only considered systematic reviews, RCT and comparative cohort studies for inclusion. In order to answer this review question, only studies assessing women who initiated HRT use before the age or average age of 65 years were included for consideration in this review.

For full details see review protocol in Appendix D.

10.3.4 Description of included studies

Four studies, of which one was a parallel RCT (Bonds 2006) and 3 were comparative cohort studies (de Lauzon-Guillain 2009; Manson 1992; Zhang 2002 were included for this review question. Although Bonds 2006 study and Manson 1992 were both WHI-related publications and some of these women were double counted in both studies, results are presented.
separately as subgroup analyses on duration of HRT are only provided in the publication by Manson 1992 that was considered as a cohort study.

Women included in 3 studies were postmenopausal (Bonds 2006, Manson 1992 and Zhang 2002) and menopausal women with no further description in one study (de Lauzon-Guillain, 2009). Participants in all studies were not diagnosed with T2DM at the baseline.

Self-reported HRT use at baseline or during follow-up either elicited by survey questionnaire (de Lauzon-Guillain 2009; Manson 1992) or ascertained by prescriptions brought to the study visit (Zhang 2002) was examined across the 3 cohort studies. Risk of T2D mellitus (T2DM) in relation to the characteristics of HRT such as user category, formulation, duration, and age was assessed across the 3 cohort studies, while the single RCT (Bonds 2006) examined the risk of T2DM associated with conjugated equine oestrogen compared with placebo. Follow-up time of the 3 cohort studies (de Lauzon-Guillain 2009; Manson 1992: Zhang 2002) ranged from an average of 4 years to 14 years, whereas the RCT (Bonds 2006 WHI) lasted for on average 7 years.

Three studies were undertaken in the USA (Bonds 2006 WHI; Manson 1992; Zhang 2002), and one in France (de Lauzon-Guillain 2009). The RCT (Bonds 2006) included 9,712 participants and the sample size of the 3 cohort studies ranged from 857 (Zhang 2002) to 63,624 (de Lauzon-Guillain 2009). The majority of studies included women with an age profile between 48 to 59 years whereas one study (Zhang 2002), which was conducted among American Indian Woman recruited women with a wider age profile, included women between 45 and 74 years.

Results from studies that did not specify the HRT type and those reporting results for the comparison of combined equine oestrogen with placebo are presented separately.

A summary of the baseline characteristics of included studies are presented in Table 20.

### Table 20: Summary characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonds 2006 (WHI data)</td>
<td>Combined equine oestrogen (current, past users) Placebo</td>
<td>Post- menopausal women with no diagnosis of diabetes mellitus at baseline and who had undergone hysterectomy N=9712 Age range: 50-59 BMI (kg/m2), n (%), p value -CEO: &lt;25: 1,073 (22.4) 25-30: 1,677 (35.1) &gt;30: 2,032 (42.5) -Placebo &lt;25: 1,046 (21.5) 25-30: 1,749 (39.9) &gt;30: 2,079 (42.7)</td>
<td>Diabetes risk (only for those aged between 50-59 years)</td>
<td>Follow-up was 14 years Exclusions: women with previous history of breast cancer, or any cancer within previous 10 years, current use of corticosteroids, anticoagulants, tamoxifen, or other selective oestrogen receptor modifiers and triglycerides &gt;4.56 mmol/l, history of thromboembolism.</td>
</tr>
<tr>
<td>Manson 1992 (subgroup analysis on WHI data)</td>
<td>Combined equine oestrogen (current, past users) Placebo</td>
<td>Postmenopausal women, free of diabetes mellitus, CHD and stroke diagnoses (N=21,028) Age range: 48-50 years BMI, mean (SD) Never: 24.6 (4.4) past: 24.3 (4.2) Current: 23.7 (3.7)</td>
<td>Risk of T2DM</td>
<td>Prospective cohort study, follow-up was 12 years Exclusions: women reporting diabetes diagnosis before 1976, women with T1DM, women with ketonuria (more than trace) on at least two occasions or hospitalisation for ketoacidosis, women classified as having gestational diabetes</td>
</tr>
</tbody>
</table>
### Table

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Lauzon-Guillain 2009</td>
<td>Menopausal hormone therapy (MHT) (oestrogen) (current and past use)</td>
<td>Menopausal women (N=63, 824) Age range: 40-65 years BMI (Kg/m2 ), mean (SD) By MHT use: --Non-user: 23.8 (3.8) -User: 22.9 (3.1)</td>
<td>Risk of T2DM (self-reported)</td>
<td>Follow-up was 14 years Prospective cohort study of women living in France who were covered by the national insurance plan for teachers and co-workers Exclusions: women who did not respond to dietary questionnaire or had miscoding of questionnaire.</td>
</tr>
<tr>
<td>Zhang 2002</td>
<td>HRT</td>
<td>Postmenopausal women who did not have a history of diabetes, did not take diabetes medication, and had a fasting plasma glucose level of &lt;7.0 mmol/l (126 mg/dl) and a 2-h post challenge glucose level &lt;11.1 mmol/l at the baseline examination(N=857) Characteristics of population; users were more educated, had a higher hysterectomy rate, had lower American Indian heritage, gravity and parity, and were more active. Current users were younger than past or never users and had lower BMI</td>
<td>HRT use (past and never users vs current users of oestrogen) and risk of T2DM Risk of T2DM and fasting glucose &gt;= 7.0 mmol/l Risk of T2DM and 2-h glucose &gt;=11.1 mmol/l (200mg/dl) Duration (as a continuous variable) of oestrogen use and risk of T2DM (fasting glucose &gt;=7.0mmol/l) Duration of oestrogen use and risk of T2DM (2-hr glucose &gt;=11.1 mmol/l)</td>
<td>Longitudinal study (cohort), follow-up 4 years Survey carried out among volunteers from 13 Indian tribes/communities Exclusions: Women who had inconsistent information on oestrogen use at the baseline and at the second examination Data was adjusted for covariates including BMI, hysterectomy status, education, family history, American Indian heritage</td>
</tr>
</tbody>
</table>

### Abbreviations:
- BMI: body mass index
- CEO: combined equine oestrogens
- HRT: hormone replacement therapy
- MHT: menopausal hormone therapy
- SD: standard deviation
- T2DM: T2D mellitus

### 10.3.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F:, study evidence tables in Appendix H:, forest plots in Appendix J:, and exclusion list in Appendix G.

### 10.3.6 Economic evidence

No health economic search was undertaken for this guideline as the decision was made to prioritise short term treatment.

### 10.3.7 Evidence statements

#### Evidence from RCTs

Low quality RCT evidence from almost ten thousand women aged 50-59 showed that there was no significant difference on the risk of T2DM between those who were current users of conjugated equine oestrogen compared to placebo at 7 years follow-up.
Evidence from cohort studies

Low to very low quality evidence from two separate cohort studies (with sample sizes of 21,028 and 63,624) showed that current HRT users have a significantly lower risk of developing T2DM compared to non-users at 12 and 14 years follow-up respectively. In addition, subgroup analysis on the different durations of treatment with HRT (less than 1 or 2 year, less than 5 years, or more than 5 or 7 years) treatment showed largely results on the same direction (very low quality evidence) although results should be interpreted with caution given the post hoc subgroup analyses of these observational studies.

Very low quality evidence from two separate cohort studies and their post hoc subgroup analyses on different durations of HRT found no significant difference in the risk of T2DM between past HRT users and non-users.

Very low quality evidence on a post hoc subgroup analyses of the route of HRT administration found that the protective effect of HRT use on the risk of T2DM was preserved either HRT was administered orally or transdermally (cohort study of over 20000 postmenopausal women).

10.3.8 Evidence to recommendations

10.3.8.1 Relative value placed on the outcomes considered

The GDG decided that T2DM and mortality (either general or condition specific) are the most important outcomes for this question. However, the GDG discussed that T2D may be unrecognised and this was taken into consideration at the time of developing recommendations.

10.3.8.2 Consideration of clinical benefits and harms

Although evidence from randomised studies showed no significant difference in the T2DM risk between HRT users and placebo, evidence from large cohort studies found that current HRT users have a significantly lower risk of T2DM compared to non-users. This protective effect of HRT on the risk of developing T2DM seems to disappear when the HRT treatment stops as it was found when results were compared between past HRT users and non-users. Results on post hoc subgroup analyses on the effect of different durations of HRT on the risk of T2DM showed that the majority of results remained in the direction of HRT being a protective influence on T2DM risk. Route of administration also did not seem to change the HRT's protective effect on T2DM. The group discussed that this result is in contrast to the combined oral contraceptive which contains higher concentrations of more potent sex steroids.

Most of the women included in the studies were postmenopausal before the age of 65 years. Although the outcome of diabetes was self-reported in most of the studies and biochemical confirmation was not necessarily obtained, the results might underestimate the protective effect of HRT on the risk of T2DM given that some cases would be undiagnosed. Only one study used the diagnosis of diabetes based on measurement of plasma glucose levels.

10.3.8.3 Consideration of economic benefits and harms

The GDG believe the clinical review revealed some evidence that HRT offered a protective effect against T2D which, depending on the magnitude of the effect, potentially could save future health service costs in the treatment and management of type 2 and its complications, as well as averting losses in HRQoL.
10.3.8.4 Quality of evidence

The evidence basis for these recommendations was one RCT and 3 comparative cohort studies. One of the cohort studies included was a post hoc subgroup analysis for a part of the same dataset that was used in the RCT. However results are presented separately due to the additional information given in the cohort for some predefined subgroup analyses in our protocol. All the subgroup analyses presented by the cohort studies should be interpreted with caution due to the risk of type II errors.

The main reasons for downgrading the quality of the studies were the high and very high risk of bias due to selection, performance and attrition bias. Quality of evidence was also downgraded due to imprecision in the estimates of relative effects.

10.3.8.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

This section refers only to menopause women with no prior diagnosis of T2D, or with insulin-dependent (type 1) diabetes.

Women with ketonuria (more than trace) were also outside the scope of this review question. There is another section in the guideline that refers to women with T2DM and the associated risk of glucose control with HRT treatment (Cross refer to 10.4).

10.3.8.6 Key conclusions

The GDG concluded that HRT administration is associated with a lower risk of developing T2DM.

10.3.9 Recommendations

44. Explain to women that taking HRT (either orally or transdermally) is not associated with an increased risk of developing type 2 diabetes.

10.4 T2D management – control of blood sugar

10.4.1 Review question

What impact does administration of HRT have on diabetes/glycaemic levels in those with T2D?

10.4.2 Introduction to topic

Diabetes is a heterogeneous condition which presents as a syndrome of biochemical and clinical disturbances of which blood glucose levels have been adopted as the defining criteria. HRT is however known to affect many biochemical markers so surveillance of all these should be continued as routine.

The menopausal transition is defined as a time of irregularity in the menstrual cycle and variation in hormone levels. Changes in sex hormones can have an influence on blood sugar levels. The symptoms of flushing and night sweats can be confused by a woman with diabetes as a symptom of hypoglycaemia.

There is some evidence that oestrogens and non-androgenic progestogens do not impair glycaemic control. Current practice is to use transdermal methods of HRT delivery in women with diabetes.
There is little evidence of significant long term changes to blood sugar levels with the administration of HRT. Normal regular assessments of diabetes control should continue with blood sugar levels being more closely monitored only at the initiation of therapy.

10.4.3 Clinical introduction

The objective of this review was to assess the impact of HRT use on diabetes/glycaemic control in menopausal (including perimenopausal and postmenopausal) women with T2D mellitus (T2DM). Comparisons were presented for any type of HRT and placebo or no HRT. Subgroup analyses were only considered based on the age distribution of the included population or on the stage of menopause (peri or postmenopausal) if data were available. Given the interventional nature of this review question we only considered for inclusion systematic reviews of RCTs, RCTs and comparative cohort studies. For full details see review protocol in Appendix D.

10.4.4 Description of included studies

Five RCTs, 4 of which were parallel RCTs (Darko 2001, Kernohan 2007, McKenzie 2003, Perera, 2001) and one crossover RCT (Sutherland 2001) were included for this review question. No comparative cohort studies were found to match our protocol. However, we identified one large (over 15,000 women) cross-sectional study (Ferrara 2001) from a USA Diabetes Register which compared different types of HRT with placebo. After discussion with the GDG it was decided that this study would be included in the review to provide supplementary evidence given its large size. Results from this study were interpreted with caution due to the limitations of this study design and the lack of confidence in the production of effect sizes.

Women included in all studies were postmenopausal women with T2D. Some common exclusion criteria were reported across these studies such as women taking insulin, lipid lowering therapy, HRT use prior to study entry, poor glycaemic control, other co-morbidities (such as breast cancer or endometrial cancer), or moderate to severe hypertension. The age of the population ranged from 60 to 70 years. The majority of studies were conducted in the UK, one study in USA and one study in New Zealand.

Results are presented separately by HRT type. The two types of HRT included were sequential and continuous combined HRT (either oral or transdermal). We only found evidence on the outcomes of glycaemic control at 12 weeks and 6 months measured by either glycosylated haemoglobin (HbA1c, %) or blood glucose levels (m/mol). No evidence was found for the other outcomes specified in the protocol (health quality of life, mortality or adverse events).

Data from the cross-over RCT (Sutherland 2001) were only reported from the second arm (after wash-out to 6 months of treatment) and were presented separately.

Evidence from these are summarised in the clinical GRADE evidence (Appendix I) See also the study selection flow chart in Appendix F, forest plots in Appendix J, study evidence tables in Appendix H and exclusion list in Appendix G.

A summary of the baseline characteristics of included studies in this review are presented in Table 21.

Table 21: Summary of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Darko 2001 (UK)</td>
<td>Sequential combined oral 17-ß oestradiol 2 mg for 16 days followed by 17-ß</td>
<td>Postmenopausal women with T2D (N=33) BMI (kg/m2): Oral</td>
<td>Glycosylated haemoglobin (HbA1c) at 12 weeks Fasting plasma glucose</td>
<td>Excluded criteria: Women taking insulin or lipid lowering therapy</td>
</tr>
</tbody>
</table>
### Menopause

**Long-term benefits and risks of HRT**

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>oestradiol 2mg plus norethisterone 1mg for 12 days</td>
<td>Postmenopausal women (&gt; 1 year from last menstrual period) with T2D (N=30)</td>
<td>Glycosylated haemoglobin (HbA1c) at 6 months Blood glucose levels (mmol/l) at 6 months</td>
<td>Exclusion criteria: poor glycaemic control (HbA1c &gt;10%), severe hypoglyceridaemia (&gt;7.0mmol/l), serum creatinine &gt;120µmol/l, blood pressure &gt;160/110 mmHg, HRT use within 2 years, insulin therapy, or other standard contraindication to HRT</td>
</tr>
<tr>
<td></td>
<td>Sequential combined Transdermal 17-β oestradiol 50µg per 24h for 14 days followed by second patch releasing both 17-β oestradiol 50µg plus norethisterone 170µg per 24h for 14 days No HRT</td>
<td>Postmenopausal women (&gt; 1 year from last menstrual period)</td>
<td>Glycosylated haemoglobin (HbA1c) at 3 months of treatment Fasting glucose (mmol/l) after 3 months of treatment</td>
<td>Exclusion criteria: poor glycaemic control, severe hypertriglyceridaemia, moderate to severe hypertension, renal impairment, liver disease, vascular disease, history of breast cancer or first degree relative with breast cancer</td>
</tr>
<tr>
<td></td>
<td>oestradiol 2mg plus norethisterone 1mg for 12 days</td>
<td>Postmenopausal women with T2D (N=43)</td>
<td>Glycosylated haemoglobin (HbA1c, %) at 6 months Blood glucose levels (mmol/l) at 6 months</td>
<td>No mention on clear inclusion or exclusion criteria</td>
</tr>
<tr>
<td></td>
<td>Transdermal 17-β oestradiol 50µg per 24h for 14 days followed by second patch releasing both 17-β oestradiol 50µg plus norethisterone 170µg per 24h for 14 days No HRT</td>
<td>Postmenopausal women with T2D (N=43)</td>
<td>Glycosylated haemoglobin (HbA1c) at 6 months Blood glucose levels (mmol/l) at 6 months</td>
<td>A cross-over trial. Authors reported data from the second arm of the trial after the washout period to 6 months HRT/placebo</td>
</tr>
<tr>
<td></td>
<td>Current use of HRT (62% unopposed oestrogen, 36% opposed oestrogen, 2% progestogens alone) No HRT use</td>
<td>Postmenopausal women with T2D and HbA1c measured during the 2 year study period (N=15, 435)</td>
<td>Glycosylated haemoglobin (HbA1c) during the 2 year study period</td>
<td>Cross-sectional study of the Kaiser Permanente Diabetes Registry Cohort</td>
</tr>
</tbody>
</table>

**National Collaborating Centre for Women’s and Children’s Health**

149
### 10.4.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H; forest plots in Appendix J; and exclusion list in Appendix G.

Study quality was assessed using the GRADE methodology. RCTs were initially assigned high quality whereas prospective cohort studies as moderate quality and downgraded based on potential sources of bias.

### 10.4.6 Economic evidence

No health economic search was undertaken for this guideline as the decision was made to prioritise outcomes from short term treatment.

### 10.4.7 Evidence statements

Very low to low quality evidence from a RCT with 24 women comparing continuous combined HRT (oral or transdermal) with placebo showed that there was no significant difference on the outcome of diabetic control as measured by either HbA1c (%) or fasting glucose levels at 3 months follow-up. The same conclusion was found from low to very low quality evidence from four RCTs (of around 47 women or less) which looked at both diabetic control measurements for continuous combined HRT users at 3 and 6 months follow-up compared to non-users.

Very low quality evidence from the only included RCT with 49 women with T2D found significantly lower levels of blood glucose at 6 months for those treated with conjugated equine oestrogen alone compared to those with placebo.

Very low quality evidence from one large cross sectional study of almost fifteen thousand women with T2D showed that when results were adjusted for women's age, there was a significant difference in the decrease of HbA1c (%) during 2 years of HRT duration for those women treated with HRT use compared to those who did not.

### 10.4.8 Evidence to recommendations

#### 10.4.8.1 Relative value placed on the outcomes considered

Glycosylated haemoglobin (HbA1c, %), blood glucose concentration (m/mol), HRQoL, mortality (overall or condition specific mortality), and adverse events (specifically complications from diabetes) were considered as the most important outcomes when considering these recommendations.

#### 10.4.8.2 Consideration of clinical benefits and harms

The only evidence found was for the outcomes of HbA1c and blood glucose measurements and for postmenopausal women. Evidence from randomised participants was presented separately by HRT type. Weak evidence showed that although treatment with conjugated equine oestrogen alone may be linked with a significant decrease in blood glucose levels at 6 months for HRT users with T2D, this direction of effect was not found when the impact of

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>cohort who stated that they did not have diabetes in the survey</td>
</tr>
</tbody>
</table>

Abbreviations: HRT; hormone replacement therapy
either sequential or continuous combined HRT on diabetic control was examined (for either 3 or 6 months outcomes). No significant change in the direction of above effects was found for either oral or transdermal HRT preparation. The GDG discussed the interpretation of these results and concluded that the lack of any significant differences between the HRT and no HRT groups would be expected given the trials’ short duration (as it would take longer for any effect on blood glucose levels to be observed).

In addition, supplementary evidence from a large cross sectional study showed that HRT may have a positive impact on diabetes/glycaemic control in menopausal women taking HRT for a longer duration (2 years) as HbA1c % was significantly reduced between the comparison groups, these results should be interpreted with caution given the lack of comparability of two groups (only adjusted for age differences) and due to outcome reporting bias (given that the exact timing of outcome reporting was unclear).

Control of blood glucose is important to prevent the acute complications of ketosis and hyperglycaemia. In addition, long-term complications such as retinopathy, neuropathy, nephropathy, and CVD can be minimised if blood glucose levels are effectively controlled. Therefore the subgroup of participants with T2DM who were controlling their blood glucose while receiving treatment for menopausal symptoms was considered highly important. The group discussed that the included evidence did not suggest that HRT was contraindicated for women with T2DM, but was not strong enough to indicate a clear benefit of improving blood glucose control. However, the Group discussed extensively how other co-morbidities should be noted when considering the use of HRT for women with T2D.

10.4.8.3 Consideration of economic benefits and harms

In the absence of evidence that HRT exerts either a negative or positive impact on diabetic/glucose control for women with T2D it is not possible to state what the economic benefits and harms are, if any.

10.4.8.4 Quality of evidence

The quality of evidence included for this question was considered to be low to very low. The included trials had very small sample sizes (the largest included fifty women in total) and there were serious concerns about the risk of bias (selection, performance and attrition). Imprecision was also a quality domain commonly and negatively affected. The timing of outcomes reported (3 to 6 months) was also not long enough to allow the demonstration of an effect between the comparisons (HRT or no HRT use). Not all studies have provided information about whether the blood glucose testing was conducted under fasting conditions.

In addition, the supplementary information from the cross sectional study gave some indication of the association between HRT use and reduction of blood glucose levels.

10.4.8.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence and on GDG expert opinion.

The group discussed the difference in diagnostic performance between HbA1c and blood glucose as a measure of diabetic control. Although the use of glucose has been considered the “gold standard” for assessing diabetic control for many years, glucose testing suffers from several deficiencies which are difficult to overcome (for example, the requirement that the subject be fasting at the time the blood is drawn and the lack of sample stability). Alternatively, measurements of HbA1c which reflect chronic blood glucose values are now routinely used in monitoring glycemic control and guiding therapy. This is because HbA1c measured using this method has been associated with a significant reduction in microvascular complications.
10.4.8.6 **Key conclusions**

The GDG concluded that HRT does not exert a negative or positive impact on diabetic/glucose control for women with T2D. However, the evidence base on this topic had flaws and the generalisation of results should be interpreted with caution.

10.4.9 **Recommendations**

45. Ensure that women with type 2 diabetes and all healthcare professionals involved in their care are aware that HRT is not associated with an adverse effect on blood glucose control.

46. Consider HRT for menopausal symptoms in women with type 2 diabetes after taking comorbidities into account and seeking specialist advice if needed.

10.5 **Breast Cancer**

10.5.1 **Review question**

What are the effects of HRT administered for menopausal symptoms on risk of developing breast cancer?

10.5.2 **Introduction to topic**

Breast cancer is the most commonly occurring cancer in the UK with almost 50,000 new cases recorded in 2011 (www.cancerresearchuk.org), which represents a crude incidence rate of 155 per 100,000 women. The true incidence of breast cancer is greatest in the older population. However, breast cancer diagnosis reaches a peak around the age of 50-59 years, at approximately 500 cases per 100,000 women in the UK. This is due to an age-associated increase in breast cancer incidence and the identification of early cases by breast cancer screening programmes offered to women between the age of 50 and 69 years. The incidence in the UK is slightly higher than average for the European Union but comparable to Germany, Denmark and Sweden. Survival after breast cancer diagnosis is around 80% at 5 years and 70% at 20 years after treatment (NICE CG80).  

Female gender and age are the most important risk factors for breast cancer. Family history is also an important factor and may be related to specific genes, two of which (BRCA1 and BRCA2) are associated with a particularly high risk of developing breast cancer and for which testing may be performed to identify women at increased risk (NICE CG164). A number of other factors have been identified which appear to be associated with an increased risk of breast cancer including alcohol intake, exposure to diethylstilbestrol and radiation (including X-rays) as well as body fatness (obesity) and height. Evidence from randomised and observational studies have identified sex steroids, particularly oestrogen/progestogen combinations, as potential risk factors, while breast feeding and physical activity have been identified as likely protective factors. However, there are other inequalities in the incidence and survival from breast cancer (www.cancerresearchuk.org). Jewish women, particularly those of Ashkenazi heritage, are at much higher risk of developing breast cancer, while women of black and other ethnic minorities are generally at lower risk. South Asian women have a lower overall risk of developing breast cancer, but those diagnosed tend to be younger and living in more deprived areas. However, irrespective of ethnicity, poorer survival from breast cancer occurs in lower socioeconomic groups (NICE CG80).

Many women do not realise that they are more likely to die from CVD than from breast cancer but the latter evokes a greater emotive response since it is associated with more deaths in women around the age of 50. To help a woman assess her risk she must understand the background prevalence, her family history, and personal factors like weight.
and alcohol consumption, as well as contraceptive use and previous breast-feeding. Her own view of the importance of any risk is also important.

It is therefore necessary to support women in coming to a decision about the use of hormonal therapies for symptoms of the menopause, and provide information about mammographic screening and the need to be ‘breast aware’.

10.5.3 Clinical Introduction

The aim of this review is to investigate the risk of developing breast cancer associated with HRT for menopausal symptoms. The focus population of this review question is peri- and postmenopausal women up to the age of 65 years old. Given that the risk of developing breast cancer may be different for women at different stages of menopause (peri or post), as with age, subgroup analysis is presented where available for the stage of menopause and different age profiles of women.

Both RCTs and comparative prospective cohorts were selected for inclusion in this review. As cohort studies are prone to selection bias, only those whose analyses adjusted for the most common confounders (such as family history of breast cancer, BMI and age of menopause or first birth) were selected for inclusion.

Two outcomes were prioritised by the GDG: risk of developing breast cancer and mortality from breast cancer. The risk of breast cancer was investigated in relation to ever HRT use (which included current and past users), current or past use of HRT compared to no use, duration of HRT and the timing since discontinuing HRT. Analyses on the type of HRT will be presented separately, when available. Otherwise, results will be presented overall for the HRT and control arms and the interpretation of these will be discussed in the LETR section in relation to attributing breast cancer risk to a specific type of HRT.

For full details see review protocol in Appendix D.

10.5.4 Description of included studies

Five RCTs comparing some form of HRT with placebo were included in this review (Anderson 2004; Cherry 2014; Manson 2013; Schierbeck 2012; and Vickers 2007):

- one trial (Schierbeck 2012) which compared HRT to placebo did not give details on the type of HRT
- two studies (The WHI [Manson 2003; 2013]; Vickers 2007) compared oestrogen plus progestogen (EP) with placebo
- one study (Vickers 2007) compared oestrogen plus progestogen versus oestrogen and evaluated the risk of developing breast cancer
- two studies (Cherry 2014; The WHI [Anderson 2004, Manson 2013]) compared oestrogen only versus placebo

In relation to the setting of the studies, the majority of included trials were conducted in the USA (Anderson, 2004; and Manson, 2003, 2013), the UK (Cherry 2014), and Denmark (Schierbeck 2012). Vickers 2007 was a multi-site study with women recruited from the following countries: UK, Australia, and New Zealand. The age range of the populations was 45 to 68 years. Duration of HRT treatment ranged from 11.9 months to 14 years.

Three trials had post-intervention follow-up and reported risk estimates for breast cancer (Cherry 2014; Manson 2003, 2013; Schierbeck 2012) for both periods (intervention and post-intervention follow-up). Post-intervention follow-up ranged from 8 to 10 years. The following table (Table 22) gives a summary of the main characteristics of the included RCTs.
Table 22: Main characteristics of included RCTs

<table>
<thead>
<tr>
<th>Studies</th>
<th>Country</th>
<th>Age in years (mean or range)</th>
<th>Sample size; HRT/Control</th>
<th>HRT type</th>
<th>Duration of Intervention (post-intervention follow-up if exists)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schierbeck 2012</td>
<td>Denmark</td>
<td>45-58</td>
<td>502 / 504</td>
<td>HRT</td>
<td>10 (5.7) years</td>
</tr>
<tr>
<td>Vickers 2007</td>
<td>UK, Australia, NZ</td>
<td>50-69</td>
<td>22199 / 2189</td>
<td>OP/O</td>
<td>11.9 months</td>
</tr>
<tr>
<td>Manson 2003, 2013 (WHI)</td>
<td>USA</td>
<td>50-59</td>
<td>2837 / 2693</td>
<td>OP</td>
<td>5.6 (8.2) years</td>
</tr>
<tr>
<td>Anderson, 2004 (WHI)</td>
<td>USA</td>
<td>50-59</td>
<td>1637 / 1673</td>
<td>Oestrogen</td>
<td>6.8 years</td>
</tr>
<tr>
<td>Cherry, 2014</td>
<td>UK</td>
<td>50-59</td>
<td>162 / 134</td>
<td>Oestrogen</td>
<td>2 (10.6) years</td>
</tr>
</tbody>
</table>

Twenty three prospective cohort studies comparing use of HRT with never use of HRT were included in the review (Jernstrom, 2003; Beral, 2003; Fournier, 2005; Schuurman, 1995; Lando, 1999; Tjonneland, 2004; Ewertz, 2005; Stahlberg, 2004; Bakken, 2011; Colditz, 1992; Grodstein, 1997; Mills, 1989; Willis, 1996; Fournier, 2008; Lund, 2007; Saxena, 2010; Schairer, 2000; Stahlberg, 2005; Folsom, 1995; Bakken, 2004; Hedblad, 2002; Manjer, 2001; Sourander, 1998). The studies were conducted in the USA and several countries in Europe (UK, France, Sweden, Netherlands, Denmark, Norway, and Finland). The age of women with menopause included in these studies ranged from 33 to 64 years. The variation in the sample size and follow-up period of these studies was wide: sample size of studies ranged from 454 to 828,923 and follow-up duration ranged from 2.6 to 12.7 years.

The following table (Table 23) gives a summary of the main characteristics of included cohorts.

Table 23: Main characteristics of included cohorts

<table>
<thead>
<tr>
<th>Studies</th>
<th>Age in years (mean or range)</th>
<th>Sample size/peri or postmenopausal</th>
<th>HRT type</th>
<th>Duration in years (mean or median)</th>
<th>Confounders in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jenstrom, 2003</td>
<td>50-64</td>
<td>6586</td>
<td>CCEP, other HRTs</td>
<td>4.1</td>
<td>Age</td>
</tr>
<tr>
<td>Beral 2003 (WHI)</td>
<td>50-64</td>
<td>1084119</td>
<td>Oestrogen, oestrogen plus progestogen, tibolone</td>
<td>2.6-4.1</td>
<td>Age, time since menopause, parity/age of first birth, family history of BC, BMI, region, deprivation index</td>
</tr>
<tr>
<td>Fournier 2005</td>
<td>52.8</td>
<td>54548/ postmenopausal</td>
<td>Oestrogen, progestogen</td>
<td>5.8</td>
<td>Time since menopause, BMI, age at menopause, parity and age at first pregnancy, family history of breast disease, oral progestogen use, oral contraceptives and previous mammography</td>
</tr>
<tr>
<td>Sourander 1999</td>
<td>60</td>
<td>7944/ postmenopausal</td>
<td>Oestrogen</td>
<td>7</td>
<td>Social class, smoking, age, BMI, diabetes, hypertension, CAD</td>
</tr>
<tr>
<td>Schuurman 1995</td>
<td>55-69</td>
<td>62573</td>
<td>HRT</td>
<td>3.6</td>
<td>Age, time since menopause, age of first birth, family history of BC, education, BMI, smoking, alcohol, oral contraceptives</td>
</tr>
<tr>
<td>Folsom 1995</td>
<td>55-59</td>
<td>41070/ postmenopausal</td>
<td>HRT</td>
<td>6</td>
<td>Age, marital status, physical activity level, alcohol use, smoking, BMI, waist/hip ratio, and parity</td>
</tr>
<tr>
<td>Lando 1999</td>
<td>55.5</td>
<td>4761/postmenopausal</td>
<td>HRT</td>
<td>12.7</td>
<td>Age, time since menopause, age of first birth, family history of BC, education, BMI, type of menopause</td>
</tr>
<tr>
<td>Bakken 2004</td>
<td>45-64</td>
<td>35456/ postmenopausal</td>
<td>Oestrogen, oestrogen plus progestogen, estriol</td>
<td>≥ 5 years</td>
<td>Age, BMI, age at menarche, use of OCs, time since menopause, family history of breast cancer, mammography, parity and age at first delivery</td>
</tr>
<tr>
<td>Tjonneland 2004</td>
<td>50-64</td>
<td>23618/postmenopausal</td>
<td>Oestrogen, sequential oestrogen plus progestogen, continuous</td>
<td>4.8</td>
<td>BMI, duration of schooling, parity, number of births, age of first birth, history of BC, alcohol consumption</td>
</tr>
</tbody>
</table>
### Menopause

#### Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Studies</th>
<th>Age in years (mean or range)</th>
<th>Sample size/peri or postmenopausal</th>
<th>HRT type</th>
<th>Duration in years (mean or median)</th>
<th>Confounders in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ewertz 2005</td>
<td>40-66</td>
<td>78380</td>
<td>HRT</td>
<td>10</td>
<td>Age of first birth, number of children, calendar period</td>
</tr>
<tr>
<td>Hedblad 2002</td>
<td>53.8</td>
<td>5862/ peri or postmenopausal</td>
<td>HRT</td>
<td>9</td>
<td>Age, BMI, smoking, HRT use, age at menarche, parity, age at menopause, history of cancer other than breast cancer or endometrium, marital status, and social class</td>
</tr>
<tr>
<td>Manjer 2001</td>
<td>54</td>
<td>5862/ postmenopausal</td>
<td>HRT</td>
<td>9.8</td>
<td>Age at baseline, height, BMI, age at menarche, nulliparity, education and smoking habits</td>
</tr>
<tr>
<td>Stahlberg 2004</td>
<td>50-60</td>
<td>10894</td>
<td>Oestrogen, oestrogen plus progestrogen</td>
<td>6.3- 7.2</td>
<td>Age, benign breast disease, menopause age</td>
</tr>
<tr>
<td>Bakken 2011</td>
<td>58.1</td>
<td>133744/postmenopausal</td>
<td>Oestrogen, oestrogen plus progestrogen, tibolone, other</td>
<td>≤ 5 years; ≥ 5 years</td>
<td>Age, type of menopause, BMI, number of full term pregnancies, age at menarche, alcohol consumption</td>
</tr>
<tr>
<td>Colditz 1992</td>
<td>30-55</td>
<td>23965/postmenopausal</td>
<td>Conjugated oestrogen</td>
<td>12</td>
<td>Age at menopause, type of menopause, BMI, number of full term pregnancies, age at menarche, alcohol consumption</td>
</tr>
<tr>
<td>Grodstein 1997</td>
<td>30-55</td>
<td>23965</td>
<td>HRT</td>
<td>14</td>
<td>Age, age at menopause, type of menopause, BMI, diabetes, high blood pressure, smoking, oral contraceptive use, family history of breast cancer, parity, age at menarche</td>
</tr>
<tr>
<td>Lund 2007</td>
<td>58</td>
<td>35453/postmenopausal</td>
<td>Oestrogen, oestrogen plus progestrogen</td>
<td>7</td>
<td>Age, BMI, family history of BC, age at menarche, parity, age of first delivery</td>
</tr>
<tr>
<td>Mills 1989</td>
<td>55.4</td>
<td>60000/pre (43.7%) and postmenopausal</td>
<td>HRT</td>
<td>6</td>
<td>Age</td>
</tr>
<tr>
<td>Saxena 2010</td>
<td>Across groups mean 56-63</td>
<td>56867</td>
<td>Oestrogen, progestogen, oestrogen plus progestogen</td>
<td>9.8</td>
<td>Age, ethnicity, history of BC, BMI, smoking, alcohol consumption, mammographic screening, parity, age of full term pregnancy, age at menopause and at menarche</td>
</tr>
<tr>
<td>Schairer 2000</td>
<td>58</td>
<td>46355/postmenopausal</td>
<td>Oestrogen, oestrogen plus progestrogen</td>
<td>10.2</td>
<td>Age, age at menopause, education, mammographic screening, BMI</td>
</tr>
<tr>
<td>Stahlberg 2005</td>
<td>50-60</td>
<td>10874/postmenopausal</td>
<td>HRT</td>
<td>6-11</td>
<td>Age</td>
</tr>
<tr>
<td>Willis 1996</td>
<td>61.4</td>
<td>422,373/postmenopausal</td>
<td>Oestrogen</td>
<td>9</td>
<td>Age, ethnicity, menopausal status, smoking, age at menarche and menopause, BMI, alcohol consumption, age of first birth, history of BC, DES and oral contraceptives use</td>
</tr>
<tr>
<td>Fournier 2008</td>
<td>40-64</td>
<td>80377/postmenopausal</td>
<td>Oestrogen, oestrogen plus progestrogen</td>
<td>&lt; 2 years; 2-4 years; ≥4 years</td>
<td>Age, menopausal status, age at menarche and menopause, breastfeeding, history of BC, physical activity, previous mammography</td>
</tr>
</tbody>
</table>

**Abbreviations:** HRT – hormone replacement therapy; BMI – body mass index; BC – breast cancer; CAD – coronary artery disease; OC – oral contraceptive

### 10.5.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.
Study quality was assessed using the GRADE methodology. RCTs and comparative prospective cohort studies were appropriate study designs for addressing this question, so were initially assigned high quality and downgraded based on potential sources of bias.

10.5.6 Evidence statements

Evidence statements for RCTs

Low to very low quality evidence from 4 RCTs (with sample sizes ranging from a thousand to more than 5000 postmenopausal women) showed that the risk of breast cancer was not significantly different between those who had received hormonal replacement treatment and those who had not.

However, evidence from 3 RCTs including the post-intervention follow-up concluded mixed results:
- very low quality evidence from a RCT with more than one thousand participants found no significant difference between any HRT use and control group during the 16 year treatment and follow-up period. The same was found by another RCT examining the effect of oestrogen in comparison with placebo during its 12.6 years treatment and follow-up period (very low quality evidence). Low quality evidence from a RCT(for the subgroup of over 5000 women aged 50-59 years) found that the risk of developing breast cancer is significantly higher for women who received oestrogen plus progestogen compared to those on placebo during 13 years of treatment and follow-up but not for women on oestrogen alone.

Evidence statements for cohort studies

Type of HRT (duration not specified)

Several cohorts of over two hundred thousand postmenopausal women found that those who received oestrogen alone or oestrogen plus progestogen had a significantly higher risk of breast cancer compared with women who had no use. The evidence was of very low quality. However, very low quality evidence from 3 cohorts on progestogen only (sample size of almost two hundred thousand women) did not find a difference in the risk between those women taking progestogen only compared to the no use group.

Ever, current or past use of HRT

Very low quality evidence from 16 prospective cohort studies of over a million and two hundred thousand postmenopausal women showed that women who had ever used HRT were significantly at higher risk of developing breast cancer compared to placebo. The same harmful conclusion was found by low to very low quality evidence for current HRT use (9 cohorts of over a million women) but not when past use of HRT (9 cohorts of over a million women) was compared with never use of HRT.

Past HRT users were not significantly different for the outcomes of breast cancer and mortality from breast cancer compared to never users (low quality evidence from 2 and 4 studies of over 500,000 women).

For studies looking separately at the components of HRT, it was found that:
- the risk of breast cancer was also found to be significantly higher for ever or current users of oestrogen plus progestogen compared to those who had never used this type of HRT (low to very low quality evidence from 4 cohorts of over seventeen thousand women)
- among current, ever or past users of oestrogen alone, only current users were at a significantly higher risk of developing breast cancer compared to placebo but not when
**Menopause**

**Long-term benefits and risks of HRT**

they were ever or only past users (very low quality evidence from 5 prospective cohort studies of over 400 postmenopausal women)

For some of these cohorts which looked at incident cases of breast cancer, significantly more ever and current HRT users were found to be at higher risk compared to never users (very low quality evidence from pooled analyses of 7 and 4 studies respectively of over 500 thousand women).

**Mortality**

Mortality for breast cancer was not found to be significantly different for either current or ever HRT users compared to those who never used HRT (very low quality evidence from 3 cohorts of over seven hundred thousand women).

**Duration of HRT use**

Inconsistent results from several cohorts were found to reveal a trend regarding the impact of the duration of HRT use on the development of breast cancer. Very low quality evidence from 4 cohorts of over a hundred thousand women found that up to 2 years of HRT use significantly increased the risk of breast cancer compared to the group of women who never used HRT. No significant difference was found for the outcome of breast cancer between those women who used HRT up to 4 years and nonusers. The risk of breast cancer was shown to increase with HRT duration of 5 to 10 years (very low quality evidence of 3 studies of 70,000 women), 10 to 14 years and more than or equal to 15 years compared to no use (moderate quality evidence of one study of over ten thousand women).

For the studies which only included oestrogen as a type of HRT, 3 cohorts of one hundred and forty thousand women found that being treated with oestrogen alone for 5 or more years significantly increased the risk of breast cancer compared to no use (very low quality evidence). The same conclusion was shown from very low quality evidence from 2 cohorts (of over one hundred thousand women) for oestrogen’s duration of 15 years or more. However, no significant difference was found for the duration of 2 years, less than 5 years, 4 to 10 years, or more than 10 years when oestrogen alone was compared with no use (moderate to very low quality evidence).

The results from studies that tested the impact of oestrogen plus progestogen duration on the risk of breast cancer compared with no use of HRT consistently found that the risk of breast cancer was significantly higher when the duration of oestrogen plus progestogen was 4 years or greater (low to very low quality evidence from pooled analysis of 3 to 6 studies with sample sizes ranging from several thousands to almost a million women).

**Time since HRT stopping**

Moderate to very low quality evidence from a cohort study of over 7000 women which examined whether time elapsed since discontinuation of HRT (up to 4 years, 4 to 10 years, 10 or more years) would impact on the risk of breast cancer did not reveal a significant difference between the HRT and no use groups. The same conclusion was found from studies that only included oestrogen alone or oestrogen plus progestogen (low to moderate quality evidence from cohorts of several hundred thousand women in oestrogen alone studies and from cohorts of less or over hundred thousand women in combination of oestrogen plus progestogen studies).

**10.5.7 Health economics profile**

No health economic studies were identified for this question.
10.5.8 Evidence to recommendations

10.5.8.1 Relative value placed on the outcomes considered

The Guideline Development Group considered breast cancer (risk and incidence) and mortality from breast cancer as the most important outcomes for answering this review question. The GDG followed the principles set up at the NICE Patient Experience Guideline (CG138) regarding the presentation of information to personalise risks and benefits as far as possible. For that purpose the use of absolute risk is preferred rather than relative risk.

Provision of information on all aspects of the benefit/risk ratio of HRT regarding short and long term consequences of treatment is of paramount importance for women’s decision making regarding the choice of treatment for menopausal symptoms (see section on Long-term benefits and risks of HRT).

10.5.8.2 Consideration of clinical benefits and harms

The included evidence from both randomised and cohort studies which showed that there may be risk of developing breast cancer during treatment associated with oestrogen plus progestogen compared to no HRT use, but this risk does not seem to be the same for those women treated with oestrogen or progestogen taken alone.

More specifically, the WHI study found that postmenopausal women aged 50-59 years old treated over 7.5 years with oestrogen plus progestogen were 3 times more per 1000 (95% C.I. 0 to 7 more) to develop breast cancer compared to women on no HRT treatment. However, this higher absolute risk was not observed in the other two RCTs which included smaller sample sizes and longer follow-up periods. The cohort studies also found that the absolute risk of developing breast cancer was significantly higher in women who ever used oestrogen and progestogen compared to those who never used it (12 more per 1000 [95% C.I 2 to 31 more]), while the results from the current users of oestrogen plus progestogen compared with never users moved in the same direction (7 more per 1000 [95% C.I 6 to 8 more]).

Duration of HRT of more than 5 years may increase the risk of breast cancer but this associated risk seems to disappear after HRT is stopped.

10 more women per 1000 women (95% C.I 3 to 19 more) treated with HRT for 5 to 10 years may develop breast cancer compared to those who have never used HRT, and this absolute risk increases to 20 more per 1000 (95% C.I 8 to 38 more) for a duration of HRT of 10 to 14 years. Most of the women in the included studies started HRT when aged between 50-59 years and the GDG discussed how this would represent the majority of women starting HRT in the UK as it is unusual for women to start HRT after the age of 60 years.

For the studies which looked at specific types of HRT, results on the impact of duration of oestrogen alone on the risk of developing breast cancer showed the same pattern, although the absolute numbers were lower (for duration of treatment of more than or equal of 5 years, 4 more per 1000 (95% C.I 2 to 8 more), and for duration of treatment of 15 or more years, 2 more per 1000 [95% C.I 2 to 5 more] compared to no users). For oestrogen plus progestogen, it was shown that the risk of breast cancer may be related to treatment duration of even less than 5 years; the absolute risk of breast cancer for women treated up to 5 years would start from 5 more per 1000 (95% C.I 2 to 8 more) compared to non-users, to 9 more per 1000 (95% C.I 4 to 16 more) for those on oestrogen plus progestogen treatment for 4 to 10 years.

The evidence found for the outcome of mortality from breast cancer came only from observational studies of several thousand menopausal women which compared either current or ever HRT users to never HRT users. This evidence showed that mortality from breast cancer was not significantly different between current or ever HRT users and those women who never been treated with HRT.
The GDG discussed the importance of finding that although risk may be increased, there does not appear to be an increase in mortality from breast cancer, suggesting that HRT stimulates the development of cancer from occult lesions already present and that the natural history of the disease is not changed. The GDG discussed that there may have been different messages communicated to women in terms of their risk of dying from breast cancer and the use of HRT, therefore they drafted a recommendation addressing both health professionals and women to inform them of the evidence that HRT use does not influence the woman's risk from dying from breast cancer above her baseline risk.

The GDG considered that the decision to offer HRT for women in menopause should be individualised taking into account of personal (baseline) risk factors for breast cancer that include genetic predisposition and lifestyle factors for example diet, exercise, alcohol consumption, smoking and reproductive history.

10.5.8.3 Consideration of health benefits and resource uses

Breast cancer is expensive to manage and treat and has significant morbidity and mortality associated with it. As an adverse event arising from HRT use it is part of an overall trade-off of risks and benefits. This trade-off was assessed formally through an economic evaluation reported in details in Appendix L.

10.5.8.4 Quality of evidence

Evidence from both randomised and comparative cohort studies was considered for this review questions and evidence was presented by HRT type when data were available. The sample size of the studies ranged from two to several thousand of participants. Four out of 5 RCTs presented information on the different types of HRT (oestrogen only, oestrogen plus progestogen) versus placebo. The studies for the comparison of oestrogen plus progestogen also presented results for a post-intervention follow-up period and these results are presented separately from the randomised period. Due to the high heterogeneity of studies with follow-up periods (age profile of women, duration of follow-up) results are presented separately for each study. The main reasons for downgrading the quality of included evidence were due to high risk of bias and imprecision around the estimates of relative effect.

Twenty-three prospective cohort studies from a variety of settings were also included in this review. The quality of evidence was downgraded mainly because of the serious risk of bias and inconsistency in the results. Inconsistency was a serious to very serious problem in the pooled analysis of cohort studies due to differences in the study characteristics and the follow-up period. The cohort study results were adjusted for different confounders which may have contributed to the observed inconsistency of results. However, given that the direction of effect across the studies was consistent, it was decided to present pooled results to facilitate the GDG’s decision making. The GDG were advised that the precision of results should be interpreted with caution, as should the results assessing the different durations of HRT treatments on the risk of breast cancer that were based on multiple subgroup analyses (potential risk of type II statistical error).

10.5.8.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

The group discussed that due to improvements made in both screening and treatment, the mortality from breast cancer in the UK has fallen substantially over the last 20 years and 5 year survival has risen from 71% to 87% (1991-2011). Women in the UK can access mammography every 3 years from the age of 50 years, which is an important way of detecting early breast cancer. There is evidence that HRT, particularly when combined
Menopause
Long-term benefits and risks of HRT

Oestrogen and progestogen, increases the density of the breast tissue and makes the
detection of small tumours more difficult. One of the outcomes of this is that women are more
likely to be recalled for further evaluation with repeat mammography, ultrasound and even
biopsy.

10.5.8.6 Key conclusions

HRT may be associated with an increased risk of breast cancer. Any increased risk of breast
cancer associated with HRT is low and should be taken in the context of the overall benefit
and risk ratio in using HRT for treating menopausal symptoms. In addition, this risk seems to
be lost when HRT is discontinued as demonstrated in the studies on the low risk of breast
cancer for past HRT users.

10.5.9 Recommendations

47. Ensure that menopausal women and healthcare professionals involved in their
care understand that HRT does not affect the risk of dying from breast cancer.

48. Using table 3, explain to women around the age of natural menopause that:

- the baseline risk of breast cancer for women around menopausal age in
  the UK varies from one woman to another
- HRT with oestrogen alone is associated with little or no increase in the
  risk of breast cancer
- HRT with oestrogen and progestogen can be associated with an
  increase in the risk of breast cancer
- any increase in risk of breast cancer is related to treatment duration and
  reduces after stopping HRT.

Table 3: Absolute rates of breast cancer for different types of HRT compared with no
HRT (or placebo), different duration of HRT use and time since stopping HRT
for menopausal women

|                                | Difference in breast cancer incidence per 1000 menopausal women (baseline
|                                | risk in the UK population over 7.5 years: 9.45 women per 1000 [ONS data,
<table>
<thead>
<tr>
<th></th>
<th>2010])</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Past users</td>
</tr>
<tr>
<td>Women on oestrogen alone</td>
<td>RCT estimate¹</td>
</tr>
<tr>
<td></td>
<td>Observational estimate</td>
</tr>
<tr>
<td>Women on oestrogen plus progestogen</td>
<td>RCT estimate¹</td>
</tr>
<tr>
<td></td>
<td>Observational estimate</td>
</tr>
<tr>
<td>Women on any HRT</td>
<td>RCT estimate</td>
</tr>
<tr>
<td></td>
<td>Observational estimate</td>
</tr>
</tbody>
</table>

HRT, hormone replacement therapy; RCT, randomised controlled trial
1 For women aged 50–59 years
Long-term benefits and risks of HRT

Research Recommendations

4. What is the difference in the risk of breast cancer in menopausal women on HRT with either progesterone, progestogen or selective oestrogen receptor modulators?

Why this is needed

Importance to 'patients' or the population: Fear of breast cancer deters many women from taking HRT, even in the presence of debilitating menopausal symptoms. There is a lack of evidence from randomised controlled trials directly comparing the risk of breast cancer in menopausal women on HRT with either progesterone, progestogen or selective oestrogen receptor modulators. There is a need for a national registry of women with breast cancer. Optimising the risk–benefit profile of HRT will potentially reduce morbidity and mortality from breast cancer in women who need HRT over the long term because of continuing menopausal symptoms.

Relevance to NICE guidance: High: the research is essential to inform future updates of key recommendations in the guideline. In the absence of good quality randomised prospective data it has not been possible for the current guidance to make recommendations concerning the best HRT regimens for minimising the risk of breast cancer.

Relevance to the NHS: NHS costs may rise if newer, more expensive preparations are shown to have an improved safety profile and uptake is likely to increase. This may in part be offset by improvements in quality of life and economic activity in women aged 50 to 59. Reduced long-term morbidity from breast cancer will potentially reduce the burden on NHS resources.

National priorities: This was identified as a priority area by the British Menopause Society in the recommendation paper submitted to the Department of Health as part of the consultation process initiated by the Coalition Government White Paper to modernise the National Health Service.

Current evidence base: There is a lack of RCT evidence for risk of breast cancer in women with menopause who are taking HRT about the direct comparisons of either progesterone, progestogen or selective oestrogen receptor modulators. There is a need for a national register for women with breast cancer.

Equality: Safer treatment options should improve availability of treatment for some women for whom it is currently not indicated, for example those at higher risk of breast cancer.

Feasibility: The study is feasible but would require a large prospective RCT with follow-up of 5 to 10 years in order to answer the question with any degree of certainty. Other outcomes e.g. cardiovascular could be studied concomitantly to make...
## Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Research question</th>
<th>4. What is the difference in the risk of breast cancer in menopausal women on HRT with either progesterone, progestogen or selective oestrogen receptor modulators?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other comments</td>
<td>The study more cost effective. Are there any ethical or technical issues? No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research question</th>
<th>5. What is the impact of oestradiol in combination with the levonorgestrel-secreting intra-uterine system (LNG-IUS) on the risk of breast cancer and venous thromboembolism (VTE)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why this is needed</td>
<td>The type of progestogen used in HRT influences the risk of breast cancer and VTE. Many women in the UK receive the progestogenic component of HRT by the use of an intra-uterine system (the LNG-IUS) which lasts 4 years. This is a very effective means of protecting the endometrium from the effect of unopposed stimulation by oestrogen alone and has few side-effects such as those associated with standard oral or transdermal preparations. However, the risk of breast cancer is uncertain as few data are available and the risk of VTE is unknown. If the risks were similar to those of oestradiol alone, rather than the combined HRT, then this would have significant public health impact in terms of breast cancer risk. It was not possible to consider this combination in the guideline because insufficient data of sufficient quality was available. A study should compare a standard combination of oestradiol with progestogen with a combination of transdermal oestradiol and the LNG-IUS, in order to assess changes in risk factors and event rates in women wishing to initiate HRT</td>
</tr>
<tr>
<td>Relevance to NICE guidance</td>
<td>High importance</td>
</tr>
<tr>
<td>Relevance to the NHS</td>
<td>This would allow women who are potentially at increased risk of developing BC or DVT to use combined HRT with no further increased risk for these outcomes. This would be an important health benefit. Cost effectiveness of LNG-IUS is unknown in this context, but after insertion, which can be undertaken in primary or secondary care, it requires no additional care over standard combined HRT and minor adverse effects are likely to be reduced</td>
</tr>
<tr>
<td>National priorities</td>
<td>N/A</td>
</tr>
<tr>
<td>Current evidence base</td>
<td>Virtually non-existent. There is one observational study of low quality of its impact on breast cancer risk and small studies reporting efficacy. However, no significant studies have been undertaken</td>
</tr>
<tr>
<td>Equality</td>
<td>No issues</td>
</tr>
<tr>
<td>Feasibility</td>
<td>No ethical or technical issues in relation to this research recommendation.</td>
</tr>
<tr>
<td>Other comments</td>
<td>There might be some support from the pharmaceutical industry. Recruitment is always more difficult when different treatment modalities are compared, especially one requires an invasive procedure.</td>
</tr>
</tbody>
</table>

## 10.6 Osteoporosis

### 10.6.1 Review question

What are the effects of HRT administered for menopausal symptoms on the risk of development of osteoporosis?
### 10.6.2 Introduction to topic

Osteoporosis is a skeletal disorder characterised by compromised bone strength that predisposes a woman to an increased risk of fracture, causing substantial pain, severe disability and a reduced quality of life. Fractures of the wrist, hip and vertebral fractures are the most common in people with osteoporosis; hip and vertebral fractures in particular are associated with decreased life expectancy. Approximately 80,000 hip fractures occur in the UK each year (costing almost £2 billion in hospital care alone), while a further 280,000 osteoporotic fragility fractures also occur annually.

Fragility fractures are defined as those that are associated with a fall from standing height or less. They are associated with osteoporosis and are more common in women than men at all ages. Although most osteoporotic fractures are seen after the seventh decade, fracture incidence increases in women at the menopause, coinciding with lower oestrogen levels, a decrease in BMD, and higher rates of bone turnover. As osteoporosis is a symptomless condition its management focuses on fracture prevention which includes strategies for case finding and prediction of fracture risk. A number of clinical risk factors for fragility fractures have been identified which include a previous fragility fracture, use of oral or systemic glucocorticoids, history of falls or family history of hip fracture, suspected secondary osteoporosis, low body mass index (BMI), smoking and a higher alcohol intake. The presence of these factors can act both as a prompt to consider a woman's future risk of fracture and to contribute to the estimate of risk using a fracture risk assessment tool, such as FRAX (WHO Fracture Risk Assessment Tool) or QFracture algorithm. These risk assessment tools estimate the predicted risk of major osteoporotic or hip fracture over 10 years, expressed as a percentage.

Treatment can then be targeted at the primary prevention of fractures (in women who have not previously sustained a fragility fracture) and secondary fracture prevention in cases of fragility fracture, particularly for postmenopausal women. A number of therapies are licensed for the treatment of postmenopausal osteoporosis, including bisphosphonates, strontium ranelate, raloxifene, denosumab, teriparatide and calcium with vitamin D. They generally increase BMD and decrease bone turnover (although teriparatide has a different mode of action). Clinical efficacy is assessed by their effect on reducing fracture incidence. HRT containing oestrogen was identified in early clinical trials as an agent that increases BMD and decreases bone turnover at the time of the menopause. However, it is not licensed in the UK for the treatment of osteoporosis, although the benefits of continued exposure to oestrogen from HRT at the menopause can be considered in the short term (benefits on fracture risk for the duration of therapy) and the longer term (delay to future fracture risk).

### 10.6.3 Clinical introduction

The aim of this review was to identify whether HRT use modifies the risk of developing osteoporosis. Further subgroup analyses were predefined in the protocol based on the effect of different durations of HRT treatment, age of HRT initiation, different HRT treatments and the time since treatment was discontinued.

Study designs included for this question were RCTs and comparative cohort studies. Only cohort studies which included appropriate adjustment for potential confounders (as outlined in the protocol) in their analysis were included.

Different types of fractures were prioritised by the GDG to be the focus of this review: any fracture, any osteoporotic fracture, any non-vertebral fracture, hip fracture, vertebral fracture and wrist fracture.

For full details see review protocol in Appendix D.
10.6.4 Description of included studies


Further unpublished data from the RCTs were included in the synthesis of evidence for this review taken from a published systematic review and meta-analysis which assessed the role of HRT on vertebral and non-vertebral fracture (Torgerson and Bell-Syer 2001a). This meta-analysis did not meet all the inclusion and exclusion criteria in our protocol and was not incorporated per se.

The majority of studies included in this review were conducted in the USA (n = 16), the UK (n = 4), Denmark (n = 6), Sweden (n = 2), Finland (n = 5), France (n = 1), Estonia (n = 1), Italy (n = 1). A number of studies were multicentre, including the UK, USA and Denmark (n = 2), the UK, Australia and New Zealand (n = 1), the UK and Canada (n = 1). One multicentre study was conducted at 38 different sites across Australia, Canada, Europe, South Africa and the USA.


The number of women participating in each study ranged widely from 36 (Wimalawansa 1998) to 140,582 (Yates 2004). The age profile of women included in each study either varied considerably or was a defined age range that was followed up for a long period of time. Therefore, the estimation of age dependent fracture risk was not possible with the available data. One study (Manson 2013, Jackson 2006) did carry out subgroup analysis for the risk of fracture according to the age of the participants. The majority of studies included women older than 45 or 50 years up to the age of 65 years old. Only one study (Melton III 1996) included a younger population of menopausal women at the start of the study (median 43.8 years, range 18 to 56) who had undergone bilateral oophorectomy. Table 24 gives a summary of the main characteristics of included studies.

Table 24: Summary characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aitken 1973</td>
<td>Oral oestrogen mestranol (20µg)/placebo</td>
<td>N=114 Health women who had undergone hysterectomy and bilateral oophorectomy for non-malignant</td>
<td>• any non-vertebral fracture</td>
<td>• double blind placebo controlled trial</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• outcomes were assessed annually</td>
</tr>
</tbody>
</table>
## Menopause

### Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagger 2004</td>
<td>Women who completed 2 to 3 years of treatment with HRT and then discontinued treatment were compared to those who were assigned to placebo</td>
<td>N=263&lt;br&gt;• women older than 45 years of age&lt;br&gt;• passed a natural menopause at least 6 months previously&lt;br&gt;• had normal bone mineral content or BMD</td>
<td>• vertebral fracture (short term)&lt;br&gt;• non-vertebral fracture (short term)&lt;br&gt;• any fracture (short term)</td>
<td>• cohort study; adjusted for age, baseline forearm BMC and spine BMD&lt;br&gt;• follow-up: 5, 11 and 15 years after stopping HRT</td>
</tr>
<tr>
<td>Banks 2004</td>
<td>• current HRT use (questionnaire)/nonusers&lt;br&gt;• postmenopausal women aged 50 to 69 years</td>
<td>N=138737&lt;br&gt;• fracture in current users of HRT compared with never users:&lt;br&gt;• duration of HRT use less than 1 year, 1 to 4 years, 5 to 9 years, ≥ 10 years</td>
<td>• cohort study; adjusted for age, region, socioeconomic status, time since menopause, BMI and physical activity&lt;br&gt;• follow-up: 1.9 to 3.9 years</td>
<td></td>
</tr>
<tr>
<td>Barrett-Connor 2003</td>
<td>Current or past use of HRT/ Never use of HRT</td>
<td>N=170852&lt;br&gt;• postmenopausal women aged 50 years or older&lt;br&gt;• at least 6 months postmenopausal</td>
<td>• osteoporotic fracture</td>
<td>• cohort study; adjusted for age, prior fracture, health status, maternal history of fracture and cortisone use&lt;br&gt;• follow-up: 1 year after BMD assessment</td>
</tr>
<tr>
<td>Bjarnason and Christiansen 2000</td>
<td>1 or 2mg oestradiol (daily oral) sequentially combined with 25 or 50 μg gestodene/placebo</td>
<td>N=278&lt;br&gt;• healthy women within 1 to 6 years of menopause&lt;br&gt;• with an intact uterus</td>
<td>• non-vertebral fracture</td>
<td>• RCT&lt;br&gt;• follow-up: 1 year</td>
</tr>
<tr>
<td>Cauley 2003</td>
<td>Conjugated equine oestrogen 0.625mg/daily or plus</td>
<td>N=16608&lt;br&gt;• hip fracture&lt;br&gt;• wrist fracture&lt;br&gt;• vertebral</td>
<td>• RCT with post follow-up (5.1 years for the combination</td>
<td></td>
</tr>
</tbody>
</table>

---

(unknown follow-up)

women who had taken HRT between oophorectomy and the time of review were excluded
### Menopause
#### Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>medroxyprogesterone acetate 2.5 mg/daily/Placebo</td>
<td>years&lt;br&gt;• hysterectomised and non-hysterectomised women</td>
<td>fracture&lt;br&gt;• non-vertebral fracture&lt;br&gt;• any fracture</td>
<td>oestrogen and 7.1 for the oestrogen alone versus placebo&lt;br&gt;Exclusions: use of tamoxifen, women who use postmenopausal hormones required a 3 month washout period prior to study entry</td>
</tr>
<tr>
<td>Cherry 2001</td>
<td>2mg oestradiol valerate/placebo</td>
<td>N=1017&lt;br&gt;• women aged 50 to 69 years admitted to coronary care units or general medical wards with a diagnosis of MI, in participating hospitals for the duration of the study&lt;br&gt;• discharged alive from hospital within 31 days of admission</td>
<td>any fracture</td>
<td>RCT&lt;br&gt;follow-up: 2 years</td>
</tr>
<tr>
<td>Delmas 2000</td>
<td>Oestradiol 1mg with norethisterone acetate 0.25 or 0.5mg daily/placebo&lt;br&gt;All women received a daily calcium supplement of 500mg</td>
<td>N=135&lt;br&gt;• aged 45 to 65 years with a lumbar spine BMD T score between -2 and plus2 (within 2 SD of the mean value for healthy young adult women)&lt;br&gt;• postmenopausal, as defined by cessation of menstrual bleeding for at least 1 year with oestradiol levels ≤ 30 pg/ml and FSH levels &gt; 40 IU/l</td>
<td>non-vertebral fracture</td>
<td>RCT&lt;br&gt;follow-up: 2 years</td>
</tr>
<tr>
<td>Engel 2011</td>
<td>• for past users, comparisons were made between those who had stopped within the last 5 years, and those who had stopped more than 5 years ago. For current users and previous users, duration of use was</td>
<td>N=70182&lt;br&gt;• women born between 1925 and 1950</td>
<td>osteoporotic fracture</td>
<td>cohort study;&lt;br&gt;adjusted for BMI, physical activity, age at menopause, parity, previous use of contraceptives, previous use of calcium supplements and educational level&lt;br&gt;follow-up: 16</td>
</tr>
</tbody>
</table>
### Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>considered (total use &lt; 2 years, 2 – 4.9 years and ≥ 5 years).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genant 1997</td>
<td>• 0.3, 0.625 OR 1.25 mg esterified oestrogens/ • placebo, n</td>
<td>• N=406 • naturally or surgically postmenopausal women • final menstrual period at least 6 months, and within 4 years of the start of the study</td>
<td>• fracture</td>
<td>• RCT • follow-up: 2 years</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Høidrup 1999,</td>
<td>Self-administered questionnaire was conducted with detailed questions regarding behavioural habits and other health related items</td>
<td>• N=6146 • participants in the Copenhagen City Heart Study (overall age 20 to 92) • postmenopausal women</td>
<td>• hip fracture</td>
<td>• cohort study; adjusted for age, BMI, physical activity, smoking, alcohol intake, cohabitation, marital status, school education, age at menopause and parity • follow-up: 15 years</td>
</tr>
<tr>
<td>Honkanen 2000</td>
<td>• HRT during follow-up compared to those who did not use HRT during follow-up (5-year inquiry)</td>
<td>• N=11798 • women aged 47 to 56 and resident in Kuopio Province, Finland</td>
<td>• wrist fracture</td>
<td>• cohort study; Adjusted for age, time since menopause, BMI, number of chronic health disorders and history of previous fractures • follow-up at 5 years</td>
</tr>
<tr>
<td>Hosking 1998</td>
<td>2 or 5 5mg alendronate, or open label oestrogen-progestogen (conjugated oestrogens [0.625mg daily] and medroxyprogesterone acetate [5mg daily] or as a cyclical regimen of 2mg of 167icronized oestrogen daily</td>
<td>• N=563 • aged 45 to 59 years and in good health. • postmenopausal for at least 6 months (confirmed by a high serum FSH)</td>
<td>• non-vertebral fracture</td>
<td>• RCT • follow-up: 2 years</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention/Comparison</td>
<td>Population</td>
<td>Comparisons/outcomes</td>
<td>Comments</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------</td>
<td>------------</td>
<td>----------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Komulainen 1998 | HRT (2mg oestradiol valerate day [1 to 21] and 1 mg cyproterone acetate [days 12 to 21] followed by a treatment-free interval [days 22 to 28])/placebo | • N=232  
• postmenopausal women aged 47 to 56  
• within 6 to 24 months of their last menstrual period | • non-vertebral fracture  
• wrist fracture | • RCT  
• follow-up: 5 years |
| Hundrup 2004 | Current users of HRT/never users. Past users of HRT discontinued < 5, 10 years compared to never users of HRT | • N=7082  
• female members of the Danish Nurses’ Organisation aged 45 years and over | • low-energy non-spinal fractures | • cohort study; adjusted for age, weight, height, menopausal status, BMD, previous fracture history, maternal hip fracture, smoking, calcium intake and multiple chronic health disorders; Adjusted for family history, BMI, and age at menopause  
• follow-up: 6 years |
| Huopio 2000 | HRT at baseline, compared to those not taking HRT at baseline: | • N=3068  
• women aged between 47 and 56 years residing in Kuopio Province, Eastern Finland in 1989 | • any fracture | • cohort study; adjusted for age, weight, height, menopausal status, BMD, previous fracture history, maternal hip fracture, smoking, calcium intake and multiple chronic health disorders  
• follow-up: 3 years after |
### Menopause

**Long-term benefits and risks of HRT**

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Lafferty 1994 | 0.625mg conjugated equine oestrogen for the first 25 days of each month from 1964 until 1983, after this time, women with an intact uterus also received 5mg medroxyprogesterone acetate from day 14 until day 25 of every 6th month | ● N=157  
● postmenopausal women (at least 12 months of amenorrhoea  
● aged between 43 and 60 years of age  
● for women with a previous hysterectomy, postmenopause was taken as the time of onset of hot flushes, or upon reaching 55 years of age | ● vertebral fracture  
● non-vertebral fracture  
● any fracture | ● cohort study; Adjusted for aged  
● follow-up: 12 years |
| Liu 2005 | Micronised progesterones 300mg/day/medroxyprogesterone acetate 10mg/day/norethindrone 1mg/day/micronised oestradiol 1mg/day/oestradiol 1mg/day plus medroxyprogesterone acetate 1mg/day/placebo. | ● N=132  
● healthy, postmenopausal women aged 45 to 60  
● less than 5 years from menopause, FSH level >40 IU/L, bone density T-score less than -2 on baseline BMD, normal mammogram and normal cervical smear within the past 6 months | ● vertebral or hip fractures | ● RCT  
● follow-up: 2 years |
| Lees and Stevenson 2001 | 1mg of 17β oestradiol plus 5 mg dydrogesterone from day 15 to 28/1 or 2 mg of 17β oestradiol plus 10 or 20 mg dydrogesterone from day 15 to 28 | ● N=579  
● women aged between 44 and 65 years  
● no previous hysterectomy  
● naturally postmenopausal (amenorrhoeic for at least 6 months) with serum FSH >20 IU/L in all cases | ● non-vertebral fracture | ● RCT  
● follow-up: 2 years |
| Lufkin 1992 | Oestrogen (0.1mg oestradiol daily delivered as a transdermal patch), medroxyprogesterone acetate (10mg/day orally for days 11 to 21)/Placebo | ● N=75  
● fully ambulatory, postmenopausal, white women aged 47 to 75 years of age  
● documented osteoporosis but no evidence of an associated disease or a history of use of any drug known to cause osteoporosis or to affect calcium levels | ● new vertebral fracture | ● RCT  
● follow-up: 1 year |
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manson 2013</td>
<td>combined equine oestrogen plus medroxyprogesterone acetate 2.5 mg/daily, placebo, combined equine oestrogen, placebo</td>
<td>N=16608 postmenopausal women aged 50 to 79 years, hysterectomised and non-hysterectomised women</td>
<td>• hip fracture&lt;br&gt;• vertebral fracture&lt;br&gt;• all fracture</td>
<td>cohort study for RCT data from WHI&lt;br&gt;follow-up: 6.6 years (combined equine oestrogen versus placebo)&lt;br&gt;follow-up: 8.2 years (combined equine oestrogen plus medroxyprogesterone acetate versus placebo)</td>
</tr>
<tr>
<td>Maxim 1995</td>
<td>conjugated oestrogen (at least 0.3 mg)</td>
<td>N=490 white postmenopausal women (last period at least 6 months ago, or bilateral oophorectomy), within 3 years of menopause</td>
<td>oestrogen users compared to non-users: &lt;br&gt;• wrist fracture, vertebral fracture, hip fracture</td>
<td>cohort study; Adjusted for age at menopause, BMI and smoking history&lt;br&gt;Demographic data were recorded during the baseline medical record review.&lt;br&gt;follow-up: 7.3 years</td>
</tr>
<tr>
<td>Melton III 1993</td>
<td>Ever use of oestrogen (for &gt; 3 months in total)/No HRT use.</td>
<td>N=463 women who underwent oophorectomy from 1959 to 1979 at the Mayo Clinic, premenopausal at the time of surgery</td>
<td>ever users compared to non-users and duration of treatment: &lt;br&gt;• hip fracture, vertebral fracture, wrist fracture</td>
<td>cohort study; Adjusted for age&lt;br&gt;follow-up: 15 years amongst survivors, 8.5 years amongst those who died</td>
</tr>
<tr>
<td>Middleton and Steele 2007</td>
<td>HRT use (24-48 months prior to 5 year visit)/ no HRT</td>
<td>N=400 women aged 50 to 54 years at baseline</td>
<td>any fracture</td>
<td>cohort study; Adjusted for baseline BMD&lt;br&gt;follow-up: 9 years</td>
</tr>
<tr>
<td>Mosekiilde 2000</td>
<td>Sequential combined HRT for women with a uterus (2mg oestradiol for 12 days, 2mg oestradiol plus 1mg norethisterone acetate for 10 days, then 1mg</td>
<td>N=1006 women with a uterus aged 45 to 58 years old, within 3 to 34 months since their last menstrual period, or experiencing perimenopausal symptoms combined with elevated serum</td>
<td>any fracture, vertebral fracture, hip fracture</td>
<td>RCT&lt;br&gt;follow-up: 5 years</td>
</tr>
</tbody>
</table>
### Study Intervention/Comparison

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paganini-Hill</td>
<td>oestradiol for 6 days of oestrogen only for women with a previous hysterectomy (2mg oestradiol daily) / no HRT use</td>
<td>FSH levels • hysterectomised women aged 45 to 52 years old with elevated FSH</td>
<td>• hip fracture • hip fractures</td>
<td>Cohort study; Adjusted for age • follow-up: 3 years</td>
</tr>
<tr>
<td>1991</td>
<td>any oestrogen use duration of HRT (&gt;=3, 4-14, &gt;=15 years) / no HRT use</td>
<td>• N=8600 • residents of Leisure World retirement community near Los Angeles, California</td>
<td>• wrist fracture, vertebral fracture</td>
<td>• cohort study; Adjusted for age; history of fracture, BMI, heart attack, alcohol consumption, cola intake and hysterectomy; blood pressure medication, non-prescription pain medication, smoking, exercise and attitude. • follow-up: 15 years</td>
</tr>
<tr>
<td>Paganini-Hill</td>
<td>Ever use of HRT (duration of HRT use 3-14 years) /never use of HRT.</td>
<td>• N=8850 • residents of a California retirement community</td>
<td>• hip fracture</td>
<td>• combined RCT and observational study; Adjusted for age, BMI, education, smoking, physical functioning construct, history of treated diabetes, family history of cancer, cholesterol. • follow-up: 7.1 years for (combined equine oestrogen versus placebo) • 5.5 years (combined</td>
</tr>
<tr>
<td>Prentice 2009</td>
<td>Combined equine oestrogen (0.625mg/daily) alone or plus medroxyprogesterone acetate (2.5 mg)/ placebo/no use of HRT/no prior use of HRT</td>
<td>• N=9129, combined equine oestrogen trial; N=15188, combined equine oestrogen plus medroxyprogesterone trial • women from the observational sub-cohort were required to be without a personal history of breast cancer and to have had a mammogram within 2 years prior to enrolment</td>
<td>• hip fracture</td>
<td></td>
</tr>
</tbody>
</table>
## Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
</table>
| **Randell 2002** | Past HRT use (> 5 years ago, before the baseline inquiry) or current use of HRT for at least 4.5 years/never-users of HRT | • N=7217  
• women aged 47 to 56 years residing in Kuopio Province Eastern Finland in May 1989  
• post-menopausal (≥ 6 months since last natural menstruation) | • any fracture, wrist fracture | • cohort study; Adjusted for age, time since menopause, BMI, number of chronic health disorders and history of previous fractures  
• follow-up: 5 years |
| **Ravn 1999**   | • 2.5 or 5mg oral alendronate/placebo                                                   | • N=612  
• healthy women aged 45 to 59 years  
• at least 6 months post-menopausal at baseline | • any fracture | • RCT  
• follow-up: 4 years |
| **Reid 2004**   | 60 or 150 mg/d raloxifene or 0.625mg/d conjugated equine oestrogens/placebo. All women were also given a daily supplement of 400 to 600mg of elemental calcium | • N=310  
• postmenopausal women aged 40 to 60 years  
• previous hysterectomy (no more than 15 years before the start of the study)  
• serum oestradiol < 73 pmol/L. FSH level of ≥ 40 mIU/mL  
• lumbar spine BMD between 2.5 SDs below and 2.0 SDs above the mean value for normal premenopausal women | • vertebral fracture | • RCT  
• follow-up: 3 years |
| **Tuppurainen 1995** | • past or present use of HRT  
• never use of HRT                                                                           | • N=3140  
• women aged 47 to 56 years old at baseline, residing in Kuopio Province, Eastern Finland | • in past or present users of HRT, compared to never users: fractures | • cohort study; adjusted for age  
• follow-up: 2.4 years |
| **PEPI 1996**   | Conjugated equine oestrogens (CEE) 0.625mg/day alone or plus medroxyprogesterone acetate | • N=875  
• surgically or naturally menopausal women (longer than 1 year, but less than 10 years since LMP) aged 45 to 64 | • any fracture | • RCT  
• follow-up: 3 years |
### Menopause

#### Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Comparisons/outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(MPA) 10mg/day for days 1 to 12 or CEE 0.625mg/day plus MPA 2.5mg/day CEE 0.625mg/day plus 173icronized progesterone 200mg/day for day 1 to 12/placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Veerus 2006</td>
<td>0.625mg conjugated oestrogens plus 2.5mg medroxyprogesterone or 5.0mg medroxyprogesterone one acetate (for women within 3 years of their last period) placebo instead of 2.5mg</td>
<td>N=1778, women aged 50 to 64 years old, postmenopausal</td>
<td>any fracture</td>
<td>RCT, follow-up: 2 to 5 years</td>
</tr>
<tr>
<td>Vickers 2007</td>
<td>Combined HRT (0.625mg conjugated equine oestrogens alone or plus 2.5mg or 5.0 or 10 mg medroxyprogesterone one acetate daily)/placebo</td>
<td>N=5692, postmenopausal women aged 50 to 69 (no menstrual period in the last 12 months, or had undergone hysterectomy).</td>
<td>any osteoporotic fracture, hip fracture</td>
<td>RCT, follow-up: 11.9 months</td>
</tr>
<tr>
<td>Weiss 1999</td>
<td>Transdermal 17 β oestradiol patch 0.025 mg, 0.05 mg, 0.06 mg, or 0.1 mg daily delivered in patches of 6.5, 12.5, 15 and 25 cm² respectively/placebo</td>
<td>N=175, women with a previous hysterectomy, if no previous oophorectomy: at least 45 years old and with ovarian failure, as evidenced by vasomotor symptoms for at least 1 to 5 years prior to enrolment, if previous oophorectomy: at least 40 years old, and 4 weeks to 5 years post oophorectomy</td>
<td>any non-vertebral fracture</td>
<td>RCT, follow-up: 2 years</td>
</tr>
<tr>
<td>WHI (Jackson 2006, Cauley 2003, Heiss 2008, LaCroix)</td>
<td>Women with uterus: 0.625mg conjugated equine oestrogens plus</td>
<td>N=16608, combined equine oestrogen plus medroxyprogesterone acetate trial; N=10,739 combined equine oestrogen trial</td>
<td>current use of oestrogen plus progestogen HRT; hip fracture,</td>
<td>RCT. After discontinuation of trial, participants were followed up as an</td>
</tr>
</tbody>
</table>
### Study

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011, Manson 2013)</td>
<td>2.5mg medroxyprogesterone acetate daily/ placebo</td>
</tr>
<tr>
<td></td>
<td>• oestrogen plus progesterone arm:</td>
</tr>
<tr>
<td></td>
<td>• postmenopausal women with an intact uterus, aged 50 to 79 years at randomisation</td>
</tr>
<tr>
<td></td>
<td>• oestrogen alone arm</td>
</tr>
<tr>
<td></td>
<td>Women without uterus:</td>
</tr>
<tr>
<td></td>
<td>• 0.625mg conjugated equine oestrogens daily/ placebo.</td>
</tr>
<tr>
<td></td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>• oestrogen plus progesterone arm:</td>
</tr>
<tr>
<td></td>
<td>• postmenopausal women with an intact uterus, aged 50 to 79 years at randomisation</td>
</tr>
<tr>
<td></td>
<td>• oestrogen alone arm</td>
</tr>
<tr>
<td></td>
<td>Comparisons/ outcomes</td>
</tr>
<tr>
<td></td>
<td>• wrist fracture (, any fracture hip fracture, wrist fracture, vertebral fracture, any fracture</td>
</tr>
<tr>
<td></td>
<td>Comments</td>
</tr>
<tr>
<td></td>
<td>observational cohort study, multiple publications have arisen from the same trial, therefore relevant results from; Stratified by age, prior disease and randomisation status in the WHI dietary intervention trial</td>
</tr>
<tr>
<td></td>
<td>follow-up: 7.2 years (combined equine oestrogen versus placebo) and</td>
</tr>
<tr>
<td></td>
<td>: 5.2 years (combined equine oestrogen plus medroxyprogesterone acetate versus placebo)</td>
</tr>
<tr>
<td>Wimalawansa 1998</td>
<td>Oral HRT</td>
</tr>
<tr>
<td></td>
<td>0.625mg daily and progesterone 150μg for 12 days each month)/ no treatment</td>
</tr>
<tr>
<td></td>
<td>All participants were also given a daily supplement of calcium and vitamin D.</td>
</tr>
<tr>
<td></td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>• N=36</td>
</tr>
<tr>
<td></td>
<td>• postmenopausal Caucasian women with established osteoporosis (defined as at least 1, but not more than 4, radiographically demonstrable atraumatic thoracic vertebral crush fractures and spine BMD 2.0 SD below the reference range for normal healthy women aged 35 years).</td>
</tr>
<tr>
<td></td>
<td>Comparisons/ outcomes</td>
</tr>
<tr>
<td></td>
<td>• non-vertebral fracture, vertebral fracture</td>
</tr>
<tr>
<td></td>
<td>Comments</td>
</tr>
<tr>
<td></td>
<td>• RCT</td>
</tr>
<tr>
<td></td>
<td>follow-up: 2 to 4 years</td>
</tr>
<tr>
<td>Yates 2004</td>
<td>Current/previous/ever use of HRT/never-use of HRT</td>
</tr>
<tr>
<td></td>
<td>Population</td>
</tr>
<tr>
<td></td>
<td>• N=140582</td>
</tr>
<tr>
<td></td>
<td>• postmenopausal women aged at least 50 years</td>
</tr>
<tr>
<td></td>
<td>Comparisons/ outcomes</td>
</tr>
<tr>
<td></td>
<td>• hip fracture</td>
</tr>
<tr>
<td></td>
<td>Comments</td>
</tr>
<tr>
<td></td>
<td>• cohort study; Adjusted for age, BMI, prior fracture, health status, maternal history of fracture and cortisone use</td>
</tr>
<tr>
<td></td>
<td>follow-up: 1 year</td>
</tr>
</tbody>
</table>
1 **10.6.5 Clinical evidence profile**

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.

Study quality was assessed using the GRADE methodology. Because of the nature of the outcomes assessed, which developed over time, RCTs were initially assigned high quality and downgraded based on potential sources of bias.

Different comparisons of HRT use were described in the included studies; ever HRT users (consisted of both current and/or past users) were compared to never users, current HRT users versus never users and current users versus no current use. Where relevant, this has been described in the GRADE tables. Similarly, where subgroup analysis was conducted regarding the age of participants, duration of use, time since stopping HRT this analysis has been presented.

1 **10.6.6 Economic evidence**

No search for health economic evidence was undertaken as it was thought that relevant studies would be identified in the health economic review on short-term treatments. Three evaluations (Zethraeus 2005, Ylikangas 2007, Lekander 2009a) in this review included fractures in the analysis. Further details of these studies can be found in a literature review in appendix L. All compared HRT to no therapy and found HRT to be cost-effective. However, the potential health benefits of HRT for preventing osteoporosis must be considered within the context of overall benefits and adverse consequences of HRT.

1 **10.6.7 Evidence statements**

**Evidence statements for RCTs**

Low quality evidence from 5 RCTs that enrolled over 5000 postmenopausal women showed a significantly lower risk of any fracture for women with current HRT use compared to no current use.

Moderate to very low quality evidence from several RCTs considering different types of fractures in women in menopause found a significantly lower risk for current users of HRT compared to no current users for the outcomes of non-vertebral fracture and wrist fracture (the sample size of included studies ranged from over 3500 to almost 15,000 hundred women). No significant difference was found for the outcomes of vertebral and hip fracture. This was very low quality evidence.

Subgroup analysis on the duration of HRT indicated that for HRT lasting up to 2 years, no significant difference was found for any type of fracture (and individual types) between current HRT users and no current users (very low quality evidence from either individual or up to 4five RCTs with sample sizes ranging from two hundred to over 4000 women).

However, HRT duration between 2 to 5 years showed significantly lower risk of any fracture, non-vertebral and wrist fracture (low to moderate quality evidence from 3two to 4 RCTs including over a thousand women) between women used HRT and non-users.

Further stratified analysis by HRT type showed that for current users of oestrogen plus progestogen, there was moderate to very low quality evidence that the risk of any fracture and vertebral and non-vertebral fracture is significantly lower with the current HRT group compared to no current use group (from a meta-analysis of RCTs with over two thousand women and a single RCT with over 16,000 women). Inconclusive evidence of a difference between the two comparison groups (of low to very low quality) was found for the direction of effect for hip and osteoporotic fractures.
Within the RCTs which compared the role of current oestrogen plus progestogen use on different types of fractures, subgroup analyses by women’s age distribution showed that the lower risk for any fracture from the HRT use was only significant in women aged 50 to 54 years and over 65 years (65-69 years) but not between 50-59 years old (low to very low quality evidence from single trials with sample sizes ranging between over 3two to 16,000 women). However results should be interpreted with caution given that the subgroup analysis on different age profiles is coming from different sources.

Moderate to low quality evidence from individual RCTs (with over ten thousand women) which included oestrogen alone as the HRT type showed that the risk of any fracture, hip, vertebral and wrist fracture may be significantly lower for current HRT users compared to no current users but not all results were in the same direction. Further subgroup analysis on this type of HRT showed no significant differences in the risk of any and hip fracture between current and no HRT users for women aged 50-59 years old (low to very low quality evidence). For women aged 60-69 years old current users of oestrogen alone, the risk of any fracture was found to be significantly lower when compared to non-current users (moderate quality evidence from a study of almost 5000 women) but not for the case of hip fracture (low to very low quality evidence from two RCTs in both interventional and post-interventional follow-up).

Evidence statements for comparative cohort studies

Moderate to very low quality evidence from eight prospective cohort studies (with sample size ranging from over 3a hundred to hundred thousand women) showed that the risk of any fracture, non-vertebral, vertebral and hip and wrist fracture was significantly lower for current HRT users compared to either non-current or never HRT users.

Subgroup analysis on the duration of HRT showed that:

- the lower risk of any fracture and osteoporotic fracture remained significantly independent of the HRT duration (less than 1 year, 1 to 4 years, for 5 to 9 years or over 10 years) for current HRT users compared to never users (low to very low quality evidence from single RCTs) but

- the lower risk of non-vertebral and hip fracture remained significantly lower in the current HRT use group compared to never users for those women on treatment for more than 10 years (low quality evidence from single RCTs)

Moderate to very low quality evidence from single prospective cohort studies (with sample size ranging from 500 to over 8000 women) did not produce consistent results for previous HRT users and never users in terms of the difference in the risk for different types of fractures. Subgroup analysis by HRT duration did not provide any more clarity in the direction of the results for the risk of fracture among ever HRT users and never users (all low quality evidence).

When the effect of timing of HRT stopping was examined, low to very low quality evidence showed that the risk of any type of fracture, non-vertebral, hip and osteoporotic fracture was not significantly different between previous HRT users and those who discontinued HRT less than 5 years ago compared to no HRT users (from individual cohorts with sample size ranging from over 400 to over 70,000 women).

Low to very low quality evidence from one cohort enrolling over 5000 women found that the risk of non-vertebral fracture was significantly lower for both the group of current oestrogen alone or oestrogen plus progestogen users compared to never users, whereas further analysis on the timing of HRT stopping did not show any differences in the fractures risk between these groups (low to very low quality evidence from single cohorts of several thousand women).
1 10.6.8 Evidence to recommendations

2 10.6.8.1 Relative value placed on the outcomes considered

3 The Guideline Development Group considered different types of fragility fractures (such as any fracture, vertebral and non-vertebral, hip, wrist and osteoporotic) as the most important outcomes to answer this review question. Of the 6 outcomes, the most important for the GDG decision-making was hip fracture as this is associated with the greatest health and personal cost, particularly as it has an increased mortality in the year following fracture. The GDG followed the principles set up at the Patient Experience Guideline (CG138) regarding the presentation of information to personalise risks and benefits as far as possible. For that purpose the use of absolute risk is preferred to relative risk. Information provision of all aspects of the benefit/risk ratio of HRT, regarding short and long term consequences of treatment, is of paramount importance for women’s decision-making regarding the choice of treatment for menopausal symptoms. The GDG did not consider other outcomes such as BMD and bone turnover markers which are proxy markers for the risk of fracture.

5 10.6.8.2 Consideration of clinical benefits and harms

6 Consistent evidence from both randomised and cohort studies demonstrated that the risk of any fragility fracture and non-vertebral fracture was significantly lower for women currently taking HRT (either oestrogen alone or for the combination of oestrogen plus progestogen) compared to non-users. The risk of hip fracture was also found to be significantly lower for those women on HRT treatment compared to the no treatment group but this finding was only supported by the prospective cohort studies.

7 The effect of duration of HRT use on the risk of fractures was investigated in both randomised and observational data. No change in the direction of observed protective effect of HRT on the risk of any fracture was found when different HRT durations were examined (short term duration of less than a year, up to 5 years, 5 to 10 or more than 10 years) in RCTs. However, the observational evidence on non-vertebral and hip fracture showed that the effect of HRT on lowering the risk of this type of fracture was only apparent for HRT durations more than 10 years.

8 Subgroup analysis was undertaken on the role of different age profile when the risk of osteoporosis was examined in relation to HRT use; randomised evidence did not support any differences in the direction of effects based on age.

9 Evidence from cohort studies also showed that the protective effect of HRT on the risk of fractures is not influenced by the time since stopping HRT implying that protection may be preserved after HRT is stopped.

10 The GDG concluded that the evidence was robust and showed a lower risk of fracture associated with current HRT use that persists after HRT is discontinued. The GDG discussed whether women should be given information about this conclusion as a long term consequence of HRT, to be considered in the context of benefits and risks (CVD, Section 10.2 and breast cancer, Section 46).

10 10.6.8.3 Consideration of economic benefits and harms

11 There appears to be a health benefit from long-term use of HRT in preventing fractures. However, these benefits need to be considered alongside other long term consequences of HRT use in order to determine if taking HRT in the long term is a good use of resources.
1. **10.6.8.4 Quality of evidence**

The majority of evidence from RCTs was rated moderate to low quality with imprecision the domain mainly affected in the quality assessment. All women included in the RCTs were postmenopausal and aged 40 to 65 years old.

The risk of fragility fracture in women around the age of menopause is influenced by a number of confounding factors such as the use of oral or systemic glucocorticoids, previous fractures, family history of osteoporosis, smoking, alcohol consumption and low BMI. Only prospective cohort studies which adjusted their analysis for some or all of the confounders set up at the protocol were considered for inclusion in this section. However, given that the included studies have adjusted for different confounders, the meta-analysis of cohorts was not considered appropriate given the heterogeneity of their data analyses. That decision led to several cases of production of individual effect estimates for the same comparison (HRT versus no treatment) but without distorting the conclusions in terms of producing benefit or harm.

In addition, the GDG expressed some concern about the generalisation of some of the findings for the outcome of hip fracture, as most of the cohorts contributing to this evidence were studied more than 20 years ago at a time when alternative interventions to HRT for the treatment of osteoporosis were not so widely available.

2. **10.6.8.5 Other considerations**

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

Hip fracture risk at the menopause is considered to be low. The group discussed that it may be that HRT has a longer term impact on hip fracture reduction (by deferred risk) in older age. However, that evidence was not available. The group also referred to NICE guideline on osteoporosis (CG146) when discussed the recommendations in this section.

3. **10.6.8.6 Key conclusions**

The GDG concluded that the current use of HRT treatment compared to non-users for women in menopause is associated with a significantly lower risk of fragility fracture and this lower risk is preserved when HRT is discontinued. Age and HRT duration may not produce any change in the direction of these conclusions.

4. **10.6.9 Recommendations**

49. Give women advice on bone health and discuss these issues at review appointments (see the [NICE guideline on osteoporosis](#)).

50. Using table 4, explain to women that the baseline risk of fragility fracture for women around menopausal age in the UK is low and varies from one woman to another.

51. Using table 4, explain to women that their risk of fragility fracture is decreased while taking HRT and that this benefit:
   - is maintained during treatment but decreases once treatment stops
   - may continue for longer in women who take HRT for longer.
### Table 4: Absolute rates of any fracture for HRT compared with no HRT (or placebo), different duration of HRT use and time since stopping HRT for menopausal women

<table>
<thead>
<tr>
<th>Treatment duration</th>
<th>Women on any HRT</th>
<th>Observational estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5 years</td>
<td>Women on any HRT</td>
<td>140 fewer (from 28 fewer to 218 fewer)</td>
</tr>
<tr>
<td>&gt;5 years</td>
<td>Observational estimate</td>
<td>15 fewer (from 11 fewer to 17 fewer)</td>
</tr>
<tr>
<td>&gt;5 years since stopping treatment</td>
<td>Observational estimate</td>
<td>18 fewer (from 15 fewer to 20 fewer)</td>
</tr>
<tr>
<td>&gt;5 years since stopping treatment</td>
<td>Observational estimate</td>
<td>2 more (from 19 fewer to 27 more)</td>
</tr>
</tbody>
</table>

HRT, hormone replacement therapy; RCT, randomised controlled trial

1. For women aged 50–59 years
4. Bagger 2004 (Baseline risk = 333 per 1000 women) (follow-up: 7 to 24 years)
5. Banks 2004; (Baseline risk = 15.4 per 1000 women) (follow-up: 2.8 years)
6. Randell 2002; (Baseline risk = 106 per 1000 women) (follow-up: 5 years)

### 10.7 Dementia

#### 10.7.1 Review question

What are the effects of HRT administered for menopausal symptoms on the risk of dementia?

#### 10.7.2 Introduction to topic

Dementia is not a normal part of ageing, it is caused by brain diseases. The umbrella term dementia is used to describe a set of symptoms that occur when the brain is affected by certain diseases or conditions.

Experts believe Alzheimer’s disease, which is the most common form of dementia, begins to develop in midlife, but is typically diagnosed after symptoms have progressed significantly.

Dementia often starts with episodic memory decline.

Due to the aging demographic of the UK population, and improvements in other disease morbidities, 1 in 3 over 65s will die with some form of dementia. Dementia affects somewhere between 670,000 and 820,000 people in the UK and number is increasing. About 40% of the UK population know a close friend or have a family member with dementia (www.alzheimers.org.uk). By 2040, the number of people affected is expected to double.

Symptoms such as memory loss and aggression can dramatically alter personality and cause distress to patients, their families and additional carers, leading to requirements for long term social care at home or in a residential institution.

Dementia costs the UK economy more than cancer and heart disease combined and the cost maybe as much as £23bn each year.

Further research into dementia is needed; there are few good studies that provide reliable information about whether a treatment will be successful in affecting the incidence or progression of established dementia. However, some studies have suggested that HRT might impact on this and the evidence is reviewed below.
10.7.3 Clinical introduction

The aim of this review was to determine the effect of HRT on dementia for women in the menopause. Specifically, this review question aimed to determine whether initiation or duration of HRT has a protective effect by delaying the onset of dementia. As dementia is an umbrella term which describes the symptoms that occur when the brain is affected by certain diseases or conditions, this review did not aim to investigate further the different types of dementia. Subgroup analysis was prespecified if data were available for postmenopausal and perimenopausal women and at different age ranges (under 50, 50-60, 60plus).

Both RCTs and comparative cohort studies were selected for inclusion in this review. As RCTs are the most appropriate study design for addressing the question, they were initially assessed as high quality and downgraded based on potential sources of bias. Cohort studies started as moderate quality and were then downgraded for other sources of bias if necessary.

The risk of developing dementia was examined in terms of different HRT types, duration and timing since stopping. Different measurements of dementia or reduced cognitive function were included and mortality overall or attributed to dementia were also selected as outcomes of interest.

For full details see review protocol in Appendix D.

10.7.4 Description of included studies

One RCT (Rasgon, 2014) and 11 cohort studies were identified to match this review protocol (Bove 2014, Fillenbaum 2001, Kang 2004, Kawas 1997, Khoo 2012, Mitchell 2003, Pettiti 2008, Ryan 2008, Shao 2012, Tang 1996, Whitmer 2011). The majority of included studies focused on dementia as a result of Alzheimer disease. Results from cohort studies were not meta-analysed given the differences in population, scales used to assess dementia and timing of outcomes reported.

One RCT (Rasgon 2014) compared the impact of continued use of HRT on cerebral function versus no HRT use who had previously discontinued HRT. The RCT compared those who had taken 17 β oestradiol for 10 years and then discontinued use with those who had taken combined equine oestrogen for 10 years and had continued use, with the outcome being measured at 2 years to see changes in cerebral metabolism. This trial included 64 women aged 50-65 years who have been using HRT for more than one year and were considered at elevated risk for dementia, as defined by having a first-degree relative with Alzheimer disease or personal history of major depression. Cerebral function was measured as an indication of dementia by neuroimaging techniques (PET scans). The duration of previous HRT use was comparable in the two groups; 10.5 (plus4.9) in the continued HRT group and 9.4 (plus6.2) in the discontinued HRT group and participants were followed up for two years.


In relation to setting of the studies, the majority of included cohort studies were conducted in US, with two studies conducted in France and Australia respectively. The sample size in the studies was ranged from 410 (Khoo, 2012) to 15646 (Kang, 2004) women in menopause. Duration of treatment ranged from two years to 10 years, with most of the studies reporting adjusted estimates for dementia risk or decline in cognitive function. Timing of initiation of HRT treatment was reported in 5 studies (Kang 2004, Khoo 2010, Pettiti 2008, Shao 2012, Whitmer 2011).
No evidence was identified for the outcome of mortality from either RCTs or cohort studies.

### 10.7.5 Summary of included studies

A summary of the studies that were included in this review are presented in Table 25.

#### Table 25: Summary of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population (N)</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rasgon 2014</td>
<td>Continued HRT use/ no HRT use (currently discontinued)</td>
<td>Women aged 50-65 years at the time of recruitment, at least one year current HT use and at least one year post-complete cessation of menses (N=64)</td>
<td>Cerebral metabolism change</td>
<td>RCT Follow-up 2 years</td>
</tr>
<tr>
<td>Mitchell 2003</td>
<td>Current, versus past, no HRT use, HRT duration more than 5 years</td>
<td>Postmenopausal women aged 43-84 years who previously participated in the BDES study (N=1462)</td>
<td>Risk of cognitive impairment</td>
<td>Prospective cohort study</td>
</tr>
<tr>
<td>Bove 2014</td>
<td>HRT/no HRT/ different HRT duration</td>
<td>Women previously enrolled on the Religious Orders Study and the Memory and Ageing Project who were free of known dementia and were 42-49 years age at menopause (N= 1884)</td>
<td>Neurological outcomes</td>
<td>Retrospective Cohort study</td>
</tr>
<tr>
<td>Ryan 2008</td>
<td>HRT (past or current use)/ no HRT use</td>
<td>Healthy postmenopausal women aged 65 years and over (N=996)</td>
<td>Cognitive decline</td>
<td>Cohort study</td>
</tr>
<tr>
<td>Whitmer 2011</td>
<td>HRT USE in mid-life/ HRT in late-life/ no HRT</td>
<td>Women who self-reported as being menopausal at the time of health check-up. Women were considered as mid-life HRT users if taking HRT and did not have a self-reported of endocrine diseases. Late-life HRT users were considered those with two or more prescriptions or refills of HRT during 4 years (N=5504)</td>
<td>Dementia</td>
<td>Prospective cohort study</td>
</tr>
<tr>
<td>Shao 2012</td>
<td>HRT/no HRT use</td>
<td>Menopausal women from the Cache county study with mean age at menopause around 48 years (N=5677)</td>
<td>Dementia</td>
<td>Prospective Cohort study</td>
</tr>
<tr>
<td>Fillenbaum 2001</td>
<td>Current versus past oestrogen use Continuous or intermittent oestrogen use/ No oestrogen use</td>
<td>African American women with unimpaired cognition (N=2705)</td>
<td>Cognitive impairment</td>
<td>Prospective cohort study</td>
</tr>
<tr>
<td>Kang 2004</td>
<td>Oestrogen only (non-users, past users, current users, current users of oestrogen plus progestogen), different timing of HRT use</td>
<td>Women with natural menopause or bilateral oophorectomy with mean age of 49-51 years at menopause (N=15646)</td>
<td>Cognitive performance</td>
<td>Retrospective Cohort study</td>
</tr>
<tr>
<td>Kawas 1997</td>
<td>Oestrogen user</td>
<td>Post or peri-</td>
<td>Dementia</td>
<td>Prospective cohort</td>
</tr>
</tbody>
</table>
**Menopause**

**Long-term benefits and risks of HRT**

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population (N=)</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khoo 2012</td>
<td>Oestrogen/oestrogen plus progestogen use for at least 12 months/never HRT users</td>
<td>Postmenopausal women who previously participated in the LAW study, with ages ranging from 41-79 years (N=410)</td>
<td>Cognitive decline</td>
<td>Prospective cohort study Follow-up 5 years</td>
</tr>
<tr>
<td>Petitti 2008</td>
<td>Oestrogen/oestrogen plus progestogen/ no oestrogen use/ different timing of HRT use</td>
<td>Women aged equal to or more than 75 years in 1998 who had been continuously enrolled in the health plan from 1992 to 1998 (N=2906)</td>
<td>Dementia</td>
<td>Prospective cohort study Follow-up 3 years</td>
</tr>
<tr>
<td>Tang 1996</td>
<td>Oestrogen use/no oestrogen use, different durations of HRT use</td>
<td>Postmenopausal women with no evidence of cognitive impairment at initial interview and no history of stroke or PD with mean age of 74 years (N=1124)</td>
<td>Dementia</td>
<td>Retrospective Cohort study Follow-up 1-5 years</td>
</tr>
</tbody>
</table>

1 **10.7.6 Clinical evidence profile**

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.

Study quality was assessed using the GRADE methodology. RCTs and comparative cohort studies were appropriate study designs for addressing this question.

2 **10.7.7 Economic evidence**

No health economic search was undertaken for this guideline as the decision was made to prioritise outcomes from short term treatment.

3 **10.7.8 Evidence statements**

**Evidence statements for RCTs**

Low to very low quality evidence from one RCT with 45 post-menopausal women showed that the risk of dementia (as assessed by cerebral metabolism change, medial cortical area decline or posterior cingulate decline) was not significantly different between those who had received HRT and those who had not after 2 years of follow-up.

**Evidence statements for observational studies**

Low to very quality evidence from several cohort studies (prospective or retrospective study) with sample sizes ranging from almost two thousand to over 10 thousand women in menopause showed that there was no significant difference in the risk of dementia (as assessed by cognitive impairment or decline with different scales) between those women who were current or past HRT users and no HRT users. The same finding was found when the effect of dementia was examined for different durations of HRT or timing of HRT initiation and in a long follow-up of 5 or 9 years (low to very low quality evidence).
However, very low quality evidence was found from one retrospective study of over 10 thousand women in menopause that the risk of dementia in 9 years follow-up was significantly lower for those used HRT treatment compared to those who had not used HRT. The same conclusion was found for a subgroup analysis of the same population aged below 80 years; they found that the risk of dementia was significantly lower for those who had previously used HRT with two or more prescriptions or refills during 4 years (very low quality evidence).

When the effect of different preparations of HRT (oestrogen alone, progestogen alone, oestrogen plus progesterone) was examined, no significant difference was found for any of the outcomes when the risk of dementia was compared between HRT users and no users (low to very low quality evidence).

### 10.7.9 Evidence to recommendations

#### 10.7.9.1 Relative value placed on the outcomes considered

The GDG has discussed that dementia and mortality (either general or condition specific) are the most important outcomes for this question. However, the GDG noted that mild cognitive impairment, although not the same as the major cognitive decline associated with dementia but usually preceding dementia, was considered for inclusion.

#### 10.7.9.2 Consideration of clinical benefits and harms

The only small RCT included for this topic showed that there is no significant difference in dementia as assessed by different measurements for postmenopausal women who received HRT for 2 years and those who did not. In contrast, the evidence from prospective cohort studies was not consistent mainly due to the heterogeneity of data included. The majority of studies, which looked at different ways to assess dementia, showed no significant difference in the development of dementia between those women who had current or prior use of HRT and those who had not. The cohorts that showed a significant reduction in the risk of dementia for previous HRT users were large size studies of menopausal women without comorbidities who mainly started HRT in midlife. There was also some evidence that showed that in the long term follow-up (7-9 years) the risk of dementia may be significantly lower with HRT use compared to no use, although the direction of this protective effect was not consistent across the included studies, therefore results should be interpreted with caution.

The GDG considered the spectrum of both randomised and cohort evidence and concluded that although there was no strong evidence base to support the protective or negative effect of HRT on the risk of dementia for women experiencing a ‘normal’, as opposed to premature menopause. There is some indication that there may a window of opportunity for lowering the risk of dementia with HRT use for women with specific preconditions, such as higher baseline risk if they have first line relatives with dementia, or for women who have POI.

However, the group did not feel confident to extrapolate from this evidence and their clinical experience on whether a consistent direction of HRT impact on dementia exists for women going through a normal menopause.

The group concluded that it would be important to advise women in the menopause that the evidence on HRT and the risk of dementia is yet to be firmly established either way (protective or harmful effect).

#### 10.7.9.3 Consideration of economic benefits and harms

There is no strong evidence of either a risk or benefit from HRT use on dementia and in the absence of such evidence it is not possible to conclude what the economic benefits and harms are, if any.
10.7.9.4 Quality of evidence

Both randomised and comparative cohort studies were considered appropriate to address this question. However, the randomised evidence was of low to very low quality as it included just one small RCT study of high risk of bias (selection and performance bias) which was also downgraded for imprecision. In addition, its population had a younger age profile (below 60 years) and the follow-up was too short (2 years) to allow any observation of the effect of HRT to the outcome of dementia.

The quality of the evidence on comparative cohort studies varied from low to very low quality. Meta-analysis was not considered appropriate due to the high heterogeneity of the population, methods of assessing dementia (from clinician’s consensus to assessment tools and imaging techniques). The majority of cohort studies employed very large sample sizes (up to 15,000 usually healthy women in menopause) controlled for the effect of the following confounders such as age, years of education, medical risks (diabetes, hypertension, hyperlipidaemia, or stroke), race, BMI, number of children on their analyses which gives more confidence in the direction and precision of effect sizes.

10.7.9.5 Other considerations

The recommendation was based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

The GDG have discussed the paucity of good randomised data in this area that would give more precise results on the risk or benefit of HRT on dementia.

10.7.9.6 Key conclusions

The GDG concluded that the evidence was not strong for either direction of risk or benefit for dementia when HRT is administered to naturally menopausal women commenced HRT before the age of 65 years. However, some large cohort studies have shown that the risk of dementia may be lower with the HRT use in long follow-up.

10.7.10 Recommendations

52. Explain to menopausal women that the likelihood of HRT affecting their risk of dementia is unknown.

10.7.11 Research Recommendations

<table>
<thead>
<tr>
<th>Research question</th>
<th>6. What are the effects of early HRT use on dementia?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why this is needed</td>
<td></td>
</tr>
<tr>
<td>Importance to ‘patients’ or the population</td>
<td>Concern about the prospect of dementia in older age is increasing and any beneficial effect on the future risk of dementia will be important to women who are considering using HRT. There is a need for good-quality observational studies on how early HRT use affects dementia risk in women with early natural menopause, including women with premature ovarian insufficiency.</td>
</tr>
<tr>
<td>Relevance to NICE guidance</td>
<td>Medium importance: Current NICE guidance (CG42) does not recommend HRT in dementia prevention; but this is based on the absence of evidence not evidence of harm? As two NICE guidelines now fail to show such evidence, there is a clear need to come to a more definitive conclusion.</td>
</tr>
<tr>
<td>Relevance to the NHS</td>
<td>If a benefit of HRT on dementia were found there would be great benefits to Public Health and the NHS as well as social care, depending on the size of the effect, as the burden of dementia care would be delayed and decreased for a large proportion of the older population.</td>
</tr>
</tbody>
</table>
Menopause
Long-term benefits and risks of HRT

<table>
<thead>
<tr>
<th>Research question</th>
<th>6. What are the effects of early HRT use on dementia?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current evidence base</td>
<td>Good quality observational data on how HRT use affects dementia risk in women with early natural menopause are needed. There are too few studies of adequate quality and both ecological and intervention studies are required. For intervention studies, follow-up is generally insufficient to reach clinically relevant endpoints relating to cognition and other features of dementia. Future research relating to HRT should include long term follow-up.</td>
</tr>
<tr>
<td>Equality</td>
<td>Not specific equality issues.</td>
</tr>
<tr>
<td>Feasibility</td>
<td>Timing is always a problem with dementia research and, with regard to HRT, this problem is amplified, since the delay between HRT use and cognitive decline is decades. This offers considerable technical issues, relating to long term follow and potential ethical issues in studying some patients with cognitive impairment who may lack the capacity to consent.</td>
</tr>
<tr>
<td>Other comments</td>
<td>No other comments</td>
</tr>
</tbody>
</table>

1 10.8 Loss of muscle mass (sarcopenia)

2 10.8.1 Review question

What are the effects of HRT administered for menopausal symptoms on the risk of developing sarcopenia?

5 10.8.2 Introduction to topic

Sarcopenia means loss of muscle mass and strength. It is not a disease or a syndrome but part of physiological ageing. The European Working Group on Sarcopenia in Older People has developed a clinical definition and consensus diagnostic criteria for age-related sarcopenia, using the presence of both low muscle mass and low muscle function (strength or performance).

Optimum muscle function is important: for example, hand strength is vital as it enables people to carry out their normal tasks of daily living. Loss of muscle strength contributes to the risk of falling, thus sarcopenia leads to an increased risk of fractures and other injuries.

Degenerative loss of skeletal muscle mass occurs at a rate of 0.5-1% per year after the age of 25. Decrease in muscle strength is also associated with ageing and sedentary lifestyles/lack of activity. Extreme muscle loss can be the result of decreasing anabolic stimulus (e.g. growth hormone or testosterone) and promotion of catabolic stimulus, such as pro-inflammatory cytokines. There may also be genetic influences.

Dual energy X-ray absorptiometry [DEXA] may be used to diagnose sarcopenia. This methodology is predictive of negative outcomes and it is also a method familiar to most clinicians. It can be assessed at the same time as measuring bone density.

Interventions that slow sarcopenia (such as exercise and good nutrition) are important as they enable postmenopausal women to maintain independent living.
Clinical introduction

The aim of this review was to investigate the risk of developing sarcopenia for menopausal women who had received HRT for treating menopausal symptoms. Subgroup analysis was specified in the protocol if data was available for postmenopausal and perimenopausal women and at different age ranges (under 50, 50-60, 60plus).

Both RCTs and comparative cohort studies were selected for inclusion in this review. RCTs were the most appropriate study design for addressing this question, so were initially assessed as high quality and downgraded based on potential sources of bias. Cohort studies started as moderate evidence and were then downgraded for other sources of bias if necessary.

The risk of developing sarcopenia was examined in terms of different HRT types, duration of HRT and time since discontinuation. The clinical outcomes for this study identified by the GDG were measures of sarcopenia such as muscle mass and strength.

For full details see review protocol in Appendix D.

Description of included studies

A total of 7 studies were included in the review. Six RCTs: 5 were double-blinded (Sipila, 2001; Armstrong, 1996; Kenny, 2005; Ribom, 2002; and Taaffe, 2005) and one was open-label (Skelton, 1999). Only 1 prospective comparative cohort study was included (Maddalozzo, 2004).

The setting of the included studies varied; United States, United Kingdom, Finland, Sweden, and Austria and Australia.

All the populations in the included RCTs were postmenopausal women with the majority of them aged below 65 years old. However, there was one trial (Kenny 2005) that only included healthy women in the community over 65 years old who were not treated with HRT for the 6 months prior to treatment commencement. The duration of HRT in these studies varied considerably from 6 weeks to 1 year.

The comparator included either placebo, no treatment, exercise or dietary supplements.

The only comparative prospective study included in this review did not use an adjusted analysis (which is the most appropriate type of data analysis for cohort studies in order to remove the selection bias associated with participants’ recruitment), therefore the confidence in their results was interpreted with caution.

Evidence was found for measurements of muscle strength and muscle mass so these were considered separately. Each of these measurements was involved in a task important for daily living (e.g. use of the thumb).

More details on each individual study can be found in the evidence tables (Appendix H).

Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F: study evidence tables in Appendix H: forest plots in Appendix J: and exclusion list in Appendix G.

Economic evidence

No health economic search was undertaken for this guideline as the decision was made to prioritise outcomes from short term treatment.
10.8.7 Evidence statements

10.8.7.1 Evidence statements for RCTs

A meta-analysis of 2 studies of forty women found a significant increase in the outcome of knee extension torque, and 2 studies found no significant difference in any type of measurement of change in knee extension strength, flexion strength, and or handgrip strength (low quality evidence). One study of over a hundred women found that there was a significantly higher increase in the outcome of adductor pollicis muscle strength in the HRT group compared to those who received no treatment (low quality evidence).

For the measurements of muscle mass, a meta-analysis of two RCTs with 80 women showed that there was a significantly higher score in the outcome of quadriceps muscle cross-sectional area for women on HRT compared to those who had not received HRT (low quality evidence). In the same direction, moderate quality evidence from one RCT found that women on HRT have significantly higher scores in the measurement of appendicular skeletal muscle mass compared to those in the non-treatment arm. Low quality evidence found no significant differences for the other outcomes of muscle mass between the comparison groups.

10.8.7.2 Evidence statements for prospective cohort studies

Very low quality evidence from one prospective cohort study of 126 women found no significant difference in muscle strength, as a composite outcome, between women who received HRT and those who did not.

10.8.8 Evidence to recommendations

10.8.8.1 Relative value placed on the outcomes considered

The GDG assessed sarcopenia as the age-related loss of lean muscle strength and muscle mass which in turn affects balance, gait and overall ability to perform tasks of daily living. The GDG considered the change in muscle strength (knee extension torque and strength, flexion, handgrip strength and adductor pollicis) to be the most important outcome for their decision-making. Change in muscle mass was assessed using either cross-sectional lean tissue area or appendicular skeletal muscle mass.

Loss of function relating to ageing was not considered for this review question.

10.8.8.2 Consideration of clinical benefits and harms

The question the group considered was whether HRT when administered for other menopausal symptoms had a positive benefit on muscle mass and strength which would be translated into better support for the skeleton and enhanced ability to undertake tasks of normal daily living. The evidence reviewed was not consistent in terms of producing benefit for improving skeletal support for women in menopause taking HRT. The only significant result from randomised evidence was for the outcomes of quadriceps muscle cross-sectional area and adductor pollicis for postmenopausal women on HRT compared to the group of no treatment. The importance of improved strength of adductor pollicis muscle which controls thumb movements for postmenopausal women was specifically discussed.

Although the increased appendicular skeletal muscle mass, which was also found to be significantly improved for women receiving HRT compared to those who did not, would improve a woman’s ability to move, get up from sitting to standing, and perform the basic tasks of daily living. This outcome was principally found in women over 65 years of age, limiting its clinical relevance to the whole age range of the population of interest.
The GDG also discussed the role of an integrated approach to improve skeletal support for postmenopausal women, such as dietary strategies, nutritional supplementation and physical exercise; however, preventing and treating sarcopenia was not the focus of this review, although the GDG discussed its importance in improving women’s muscle strength outcomes.

### 10.8.8.3 Consideration of economic benefits and harms

Whilst the GDG concluded that there was some weak evidence that HRT improves muscle mass and strength, the interpretation and generalisability of the results was not clear and therefore it is difficult to describe what the economic benefits and harms are, if any.

### 10.8.8.4 Quality of evidence

The quality of the randomised evidence was from moderate to very low due to high risk of bias of some of the included studies (due to unclear randomisation or allocation concealment) and to imprecision. All the included women were postmenopausal with most of them under the age of 65 years old (only one study included women over 65 years old). The duration of HRT varied between the included studies, and no data were available relating outcomes to time since discontinuation.

### 10.8.8.5 Other considerations

The recommendation was based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

The GDG focused the aim of this review question on the impact of HRT on the risk of developing sarcopenia and did not consider primary treatment for that condition. For this reason, exercise that increases muscle strength was not considered as a focus in this section. There is a separate section in the guideline that looked at the role of HRT on the outcome of osteoporosis (Section 10.6). The GDG discussed the link between bone strength and reducing fractures and falls for women in menopause.

The group noted that the extension in women’s expected lifespan and increasingly sedentary lifestyles raise a great challenge for the musculoskeletal system.

### 10.8.8.6 Key conclusions

The GDG concluded that there was some weak evidence that HRT improves muscle mass and strength however there is a limitation on the interpretation of generalisation of these results.

### 10.8.9 Recommendations

53. **Explain to women that:**

- there is limited evidence suggesting that HRT may improve muscle mass and strength
- muscle mass and strength is maintained through, and is important for, activities of daily living.
11 Premature ovarian insufficiency

Premature ovarian insufficiency (POI), previously known as premature ovarian failure, means the loss of normal ovarian function, from a variety of causes, before the age of 40 years. Roughly 1 in 100 women in the UK have POI, and often the diagnosis is extremely delayed. About 1 in 1000 women are affected under the age of 30 years (Coulam 1986).

There are 3 main identifiable causes of POI: genetic, autoimmune, or iatrogenic (Yanuz 2014):

- genetic conditions include a strong maternal family history, 45, X, 46, XX and 46, XY POI, and POI associated with galactosaemia and FMR premutations
- women with an autoimmune predisposition may develop autoimmune POI, with or without other autoimmune diseases (Diabetes mellitus, Addison's, thyroid)
- women with iatrogenic menopause form an increasingly large group. These are women whose treatments for cancer (hormonal, chemotherapy and/or radiotherapy) have brought about an earlier menopause
- in most women the cause of an early menopause is unknown

Women with untreated POI (particularly surgical menopause) are at increased risk of developing osteoporosis, CVD, dementia, and Parkinsonism: all these conditions increase an earlier mortality risk.

As well as managing clinical and physical issues, these young women need support and holistic care with a number of psychosocial issues, such as infertility, sexuality and psychological distress. The specific risk factors for POI and evidence for hormone replacement in these women are considered in this chapter.

11.1 Diagnosis of premature ovarian insufficiency

11.1.1 Review question

What is the diagnostic accuracy of the following in the diagnosis of POI: cycle irregularity, vasomotor symptoms, FSH, AMH, AFC, Inhibin B, Inhibin A, Oestrogen, Ovarian volume?

11.1.2 Clinical introduction

Whilst menstrual history (cycle irregularity or amenorrhoea) in women under the age of 40 years is often the first suggestive indication of a diagnosis of POI, confirmatory testing may be required. This review aimed to identify the diagnostic accuracy of different tests for the diagnosis of POI in women.

The aim of this review was to determine the diagnostic accuracy of cycle irregularity, vasomotor symptoms, biochemical measurements (FSH, AMH, AFC, Inhibin B, Inhibin A, oestrogen) and ultrasound features (ovarian volume) in the diagnosis of POI. These indices were considered either individually or in combination.

Certain groups of women are known to be at increased risk of POI such as women with a history of chemotherapy, certain autoimmune diseases, a family history of POI, and women with chromosomal abnormalities such as Turner Syndrome and information is presented separately for these populations if data were available.

For full details see review protocol in Appendix D.
11.1.3 Description of included studies

Three studies (Giuseppe 2007, Hagen 2010, Jadoul 2011) conducted in women at high risk of POI were included in the review and their results are presented separately due to differences in their population characteristics.

One study (Giuseppe 2007) investigated the role of FSH, AMH, inhibin B and antral follicle counts as potential markers of ovarian function for 29 women who had undergone chemotherapy for the treatment of Hodgkin’s lymphoma and they were in complete remission (less than 3 years). The mean age of participants in this study was 28.5 years (SD 7.3). POI was defined only in relation to amenorrhoea and there were no details on its duration.

The aim of the second study (Hagen 2010) was primarily to determine a reference range for AMH levels in healthy girls and adolescents and, secondary, to define the diagnostic accuracy of AMH levels in the identification of POI in women with Turner syndrome. Only the information for the diagnosis of POI in women with Turner syndrome is included for the purpose of this review. POI was defined as absent spontaneous puberty, or spontaneous puberty followed by oestrogen therapy due to a lack of pubertal progression, or secondary amenorrhoea. All participants were aged between 12 and 25 years.

The third study (Jadoul 2011) conducted in 35 high risk woman aged between 16 and 46 years who had undergone bone marrow transplantation and assessed the utility of AMH and oestradiol levels in distinguishing between those with and without ovarian insufficiency. In this study, ovarian insufficiency was defined as absent pubertal development or progression, or secondary amenorrhoea confirmed by the observation of menopausal FSH levels (without further details on the threshold of FSH levels adopted).

No studies were identified which considered the diagnostic accuracy of vasomotor symptoms, inhibin A or ovarian volume for diagnosis of POI.

11.1.4 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.

Prospective or retrospective case series were considered appropriate to answer this review question.

A summary of the findings is also presented in the following graphs for easier interpretation separately for single and combination tests (green colour demonstrates useful test, red not useful and orange neutral).
Figure 9: Single tests for diagnosis of POI in high risk women

Source: <Insert Source text here>

Figure 10: Single tests for exclusion of POI in high risk women
11.1.5 Economic evidence

A single search was undertaken for health economic evidence on the diagnosis of premature ovarian insufficiency (POI). A total of 32 articles were identified by the search. After reviewing titles and abstracts, no papers were considered suitable and no full copies of papers were obtained. Therefore, no relevant economic evidence was identified for this question.

Illustrative costs are shown for these tests in Table 26 below.

Table 26: Unit costs of tests for premature ovarian insufficiency

<table>
<thead>
<tr>
<th>Test</th>
<th>Cost</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSH</td>
<td>£11</td>
<td>GDG estimate</td>
</tr>
</tbody>
</table>
Menopause
Premature ovarian insufficiency

<table>
<thead>
<tr>
<th>Test</th>
<th>Cost</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oestradiol</td>
<td>£11</td>
<td>GDG estimate</td>
</tr>
<tr>
<td>AMH</td>
<td>£35</td>
<td>GDG estimate</td>
</tr>
<tr>
<td>Inhibin B(^a)</td>
<td>£224</td>
<td><a href="http://www.medi-labs.com/a-z-of-tests">http://www.medi-labs.com/a-z-of-tests</a> accessed May 2015(^b)</td>
</tr>
<tr>
<td>AFC</td>
<td>£52</td>
<td>NHS Reference Costs 2013-14(^c)</td>
</tr>
</tbody>
</table>

(a) This is not undertaken in clinical practice
(b) This is a private sector cost
(c) Directly accessed diagnostic services, Currency code MA36Z

11.1.6 Evidence statements

Single tests

Low to very low quality evidence from three case series investigating the diagnostic accuracy of AMH on POI found that:

- a cut-off level of < 2pmol/L gave a low sensitivity, moderate specificity, and not useful positive or negative likelihood ratio for diagnosis of POI for young women affected by Hodgkin disease and may not be useful as a diagnostic tool for this population
- a cut-off level of < 8pmol/L gave a high sensitivity, low specificity, not useful positive likelihood ratio and very useful negative likelihood ratio for women with Turner syndrome or bone marrow transplantation so the use of this tool may be useful for the diagnosis of POI these populations although results should be interpreted with caution due to wide range of the confidence intervals

Low quality evidence from one case series for women with Hodgkin disease found that the diagnostic accuracy of inhibin B with a cut-off < 60 pg/mL may be not of use to diagnose POI as it was found to a low sensitivity, moderate specificity, and not useful positive or negative likelihood ratio.

The role of oestradiol (cut-off < 50 pg/mL) to diagnose POI in a group of women treated by bone marrow transplantation was found not to be useful as it demonstrated a low sensitivity and specificity, and not useful positive or negative likelihood ratio (moderate quality evidence)

Low to very low quality evidence from two case series showed that a higher cut off point of FSH (>30 mIU/mL) may be useful to rule out women who do not experience POI and for those women whose FSH is measured prior to starting hormonal treatment:

- a cut off level of ≥ 10 mIU/mL gave a low sensitivity, moderate specificity and not a useful positive or negative likelihood ratio (women with Hodgkin’s disease)
- a cut-off level of > 30 mIU/mL gave a low sensitivity, high specificity, very useful positive likelihood ratio and not a useful negative likelihood ratio for women with bone marrow transplantation when FSH was measured before starting hormone therapy but high sensitivity and specificity, and a very useful positive and negative likelihood ratio for those who had already started therapy

Low quality evidence from one case series found a moderate sensitivity, low specificity and not useful positive or negative likelihood ratio for the diagnostic accuracy of antral follicle count to diagnose POI in a group of women with Hodgkin’s disease.

Combination tests

Very low quality evidence from one case series which included women with Hodgkin’s disease found the following results in relation to the use of combination tests to diagnosis of POI:
• the combination of FSH and AMH gave a low sensitivity, moderate specificity and not
  useful positive or negative likelihood ratio
• the combination of an antral follicle count either with inhibin B or AMH levels gave a
  moderate sensitivity and specificity and moderately useful positive and negative likelihood
  ratio

6 11.1.7 Evidence to recommendations
7 11.1.7.1 Relative value placed on the outcomes considered
The GDG has considered all the properties of diagnostic accuracy measurements required
for decision-making: sensitivity, specificity, positive and negative likelihood ratio and Area
under the Curve (AUC). The GDG considered the relative importance of having a high false
positive and high false negative result from the diagnosis of POI and consequences in
women’s further clinical management. They concluded that it is equally important to have a
correct positive diagnosis can be used to initiate the appropriate treatment (please see
following section on the management of women with POI) and a correct negative diagnosis
that will prevent women from unnecessary distress and additional pharmacological
treatment.

17 11.1.7.2 Consideration of clinical benefits and harms
There was limited evidence looking at different high risk groups for POI (Hodgkin’s disease,
treated with bone marrow transplantation, Turner syndrome) that showed that the threshold
of AMH levels of < 8.8pmol/L more may useful to diagnose POI although there is high
uncertainty around this result due to the wide range of confidence intervals. In addition, FSH
levels more than 30 mIU/mL can be a useful diagnostic tool for women at high risk of POI
who had already started hormonal therapy. The group discussed how some women present
for the diagnosis of POI to be confirmed whilst on HRT, when there may have been some
doubt about the pre-treatment diagnosis but treatment was started because of symptoms; if
despite HRT, their FSH levels are still elevated it is very likely they have POI.

The group discussed the clinical relevance of these results and the limitations of the
interpretation of FSH levels as they tend to fluctuate widely over time. Therefore the repeat of
FSH measurements in two blood tests between 4-6 weeks apart was considered the most
appropriate strategy to assess FSH levels for women with POI. The selection of time interval
(4-6 weeks) was based on Committee’s expert opinion following standard clinical practice in
order to best capture any fluctuations of FSH in a period around menstrual cycle. The GDG
also considered the serum level of anti-Mullerian hormone (AMH) as a diagnostic tool for POI
as AMH does not fluctuate significantly within the menstrual cycle or between cycles.
However, AMH may produce inconsistent results depending on the AMH assay used. AMH
levels may also be affected by the use of an oral contraceptive pill and hormone therapy so
the group concluded that it should not be used in isolation to diagnose women with POI. The
group discussed that there is good evidence for use of AMH in managing fertility treatments,
but there is limited evidence for diagnosis of POI based on this tool.

The review of evidence on the other tools such as inhibin B, oestradiol and antral follicle
count and the combination of tests found that these tests may not be useful in the diagnosis
of POI. The GDG also discussed how these tests are not used routinely in current practice to
make the diagnosis of POI and they did not feel that this practice should be reviewed in the
context of developing recommendations for this topic.

45 11.1.7.3 Consideration of economic benefits and harms
The GDG noted that the AMH assay is expensive and not generally available in primary or
secondary care and cannot be justified for routine use for diagnosing POI with existing
evidence on its diagnostic accuracy. The measurement of FSH is widely available and inexpensive to perform.

However, the GDG considered the improved diagnostic accuracy from elevated FSH levels on 2 blood samples taken 4–6 weeks apart justified the costs of an additional test. Importantly, the GDG considered that the woman’s clinical and family history was important in making a diagnosis and that, as part of standard practice, would incur negligible additional opportunity costs.

8 11.1.7.4 Quality of evidence

The majority of evidence was low to very low quality as the included studies (case series) were small and at serious risk of bias. Measurements of sensitivity and specificity requires a clinically relevant threshold to be defined but the evidence was presented based on the thresholds selected by the authors. There was a high variability between the included studies in the selection of the population, definition of diagnostic tools and the measurements reported but this is not unusual for diagnostic studies.

15 11.1.7.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion.

The GDG discussed that increased awareness that irregular periods may be due to POI is necessary among health professionals and women and there may be challenges with definite diagnosis of POI.

The diagnosis of POI has profound short and long term implications for young women; it is vital that the diagnosis is made only when there is sufficient certainty from the clinical and hormonal findings.

If there is doubt about the diagnosis of POI, the woman should remain under surveillance until the diagnosis is confirmed or fully excluded.

The GDG noted that specialist societies recommend that women with POI should be kept under long-term surveillance.

11.1.7.6 Key conclusions

The GDG concluded that diagnosis of POI should be based on both assessing women’s clinical history and elevated FSH levels. Among the other diagnostic tests reviewed in this section, the GDG concluded only if there is a doubt about definite diagnosis of POI, the use of AMH can be considered.

11.1.8 Recommendations

54. Take into account the woman’s clinical history (for example, previous medical or surgical treatment) and family history when diagnosing premature ovarian insufficiency.

55. Diagnose premature ovarian insufficiency in women aged under 40 years based on:
   - menopausal symptoms, including no or infrequent periods (taking into account whether the woman has a uterus) and
   - elevated FSH levels on 2 blood samples taken 4–6 weeks apart.
56. Do not diagnose premature ovarian insufficiency on the basis of a single blood test.

57. Do not routinely use anti-Müllerian hormone testing to diagnose premature ovarian insufficiency.

58. If there is doubt about the diagnosis of premature ovarian insufficiency, consider anti-Müllerian hormone testing after seeking specialist advice (see the NICE guideline on fertility).

11.2 Management of premature ovarian insufficiency

11.2.1 Introduction to topic

Women with POI are oestrogen-deficient and treated with HRT up to the age that they would normally expect a ‘natural’ menopause, around the age of 50 years, providing there are no contra-indications to hormone therapy. Some women are treated with conventional HRT, others take the combined oral contraceptive pill. The Pill provides contraceptive cover, if that is required (as 5-10% of women with POI still conceive spontaneously), whereas conventional HRT is not a contraceptive. The Pill may be seen as socially acceptable for young women and avoids the stigma associated with premature menopause. It is also free of prescription charges. However, HRT provides physiological replacement of hormones and may be a better option for sustaining long term health. This section looks at the evidence-based advantages and disadvantages of both treatments for women with POI.

11.2.2 Review question

What is the effectiveness of HRT compared with combined oral contraceptives for the management of POI?

11.2.3 Clinical introduction

The purpose of this review was to compare the clinical effectiveness of combined oral contraceptives (OCP) with HRT for women with POI in order to determine the best way of replacing oestrogen in women with POI. The focus population of this review question was women below the age of 40 years of age with POI for any reason, including women with Turner syndrome (TS) as 90% of TS women have primary amenorrhoea. HRQoL, markers of bone density, markers of cardiovascular/metabolic health, menopausal symptoms, sexual function, adverse effects and treatment discontinuation were considered outcomes of interest.

For full details see review protocol in Appendix D.

11.2.4 Description of included studies

Two studies were identified which met the inclusion criteria for this question (Guttmann 2001, Langrish 2009). Both studies were randomised, open-label controlled trials with a cross-over design. The included studies were conducted in United Kingdom and Israel.

The first study involved 17 women with Turner syndrome compared the short term effects of HRT (0.625mg conjugated oestrogen continuously combined with 5mg medroxyprogesterone acetate for 14 days per month) with an OCP (30μg ethinyl oestradiol and 75μg gestodene). After a 4-6 month washout period, women were randomly assigned to one of the 2 interventions for 6 months, followed by the alternate intervention for a further 6 months. Effects were measured during the last month of treatment for each preparation.
Outcomes reported measures of cardiovascular health (HDL and LDL cholesterol, triglycerides), markers of bone turnover (Vitamin D metabolites, urinary deoxypyridinoline cross-links, osteocalcin and alkaline phosphatase).

The second study by Langrish recruited a total of 42 women with POI of different aetiologies, including Turner syndrome, surgical, idiopathic and post-chemo or radiotherapy. After a 2 month washout period, women were randomised to treatment with either “physiological sex steroid replacement” (defined as an HRT preparation comprising transdermal oestrogen and either vaginal or oral progestogens) or “standard hormone replacement” (defined as an oral contraceptive pill). Treatment was continued for a total of 12 months. A further 2 month washout period was then conducted before participants were switched to the alternative treatment for the remaining 12 months. Due to a number of withdrawals during the first washout period, only 34 women were eventually randomised to receive treatment. The outcomes reported were measures of BMD, indicators of cardiovascular health, and discontinuation rates (total discontinuation rates, and discontinuation due to adverse effects).

No data were identified regarding HRQoL, changes in menopausal symptoms, adverse effects (not precipitating withdrawal from the trial), cancer incidence or sexual function.

More details on each individual study can be found in the evidence tables.

### 11.2.4.1 Summary of included studies

A summary of the studies that were included in this review are presented in Table 27.

#### Table 27: Summary of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention/Comparison</th>
<th>Population</th>
<th>Outcomes</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guttmann 2001</td>
<td>Sequential conjugated oestrogen (0.625mg) for 14 days, followed by conjugated oestrogen (0.625mg) plus medroxyprogesterone acetate (5 mg) for 14 days Ethinyloestradiol (30µg) plus gestodene (75µg) for 6 months</td>
<td>N=17 Women with Turner Syndrome who were otherwise healthy</td>
<td>Fasting glucose (mmol/l) Insulin (mml/l) Triglyceride (mmol/l) Cholesterol (mmol/l): ALP (U/l) (mean, SD): 250HD (µg/l) (mean, SD): 1, 25 (OH)2 D3 (ng/l) (Osteocalcin (µg/l)) Deoxypyridinoline (µmol/mol) Endometrial thickness (mm) Uterine pulsatility index (mean, SD):</td>
<td>RCT with cross-over design with 4-6 months washout period followed by 6 months treatment Study was not blinded and no washout period was conducted between trial interventions. No analysis was conducted to assess treatment order effect Follow-up: 12 months</td>
</tr>
<tr>
<td>Langrish 2009</td>
<td>Transdermal oestradiol (100µg) daily for one week, followed by transdermal oestradiol (150µg) for 2 to four weeks) combined with 200mg progesterone pessaries twice daily in weeks 3 to 4 (some women used oral progesterone in preference to vaginal pessaries: dydrogesterone 10mg twice daily) Ethinyloestradiol (30µg) and norethisterone (1.5mg) daily for weeks one to 3, followed by 7 pill-free days</td>
<td>N=42 Women with POI attributed to chemotherapy or radiotherapy, idiopathic or surgical treatment of Turner Syndrome</td>
<td>Blood pressure and arterial stiffness at 12 months (mean difference, 95% confidence intervals): Mean difference in systolic blood Mean difference in diastolic blood BMI was unchanged throughout study Discontinuation rate BMD Mean difference in lumbar spine BMD z-score BMD outcomes: Bone ALP and PINP response to OCP Endometrial thickness Uterine volume HRT Uterine artery resistance index Uterine artery pulsatility index</td>
<td>Open label randomised, controlled cross-over trial Follow-up was one year for intervention and comparator treatments</td>
</tr>
</tbody>
</table>
11.2.5 Clinical evidence profile

Evidence from these studies is summarised in the clinical GRADE evidence profiles (Appendix I). See also the study selection flow chart in Appendix F, study evidence tables in Appendix H, forest plots in Appendix J, and exclusion list in Appendix G.

RCTs were selected for inclusion in this review question. RCTs were initially assessed as high quality and downgraded based on potential sources of bias.

11.2.6 Economic evidence

No health economic search was undertaken for this guideline as the intervention and comparator are both relatively low cost and because it was thought a priori that they would be similarly effective.

Some illustrative costs of the treatments are indicated in Table 28 below.

<table>
<thead>
<tr>
<th>Table 28: Treatment costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
</tr>
<tr>
<td>Oestrogen only oral(^a)</td>
</tr>
<tr>
<td>Oestrogen only patch(^b)</td>
</tr>
<tr>
<td>Oestrogen and progestogen oral(^a)</td>
</tr>
<tr>
<td>Oestrogen and progestogen patch(^b)</td>
</tr>
<tr>
<td>Combined oral contraceptive</td>
</tr>
<tr>
<td>Combined oral contraceptive</td>
</tr>
<tr>
<td>Combined oral contraceptive</td>
</tr>
</tbody>
</table>

\(^{(a)}\) 3 months supply
\(^{(b)}\) 1 months supply
11.2.7 Evidence statements

Low quality evidence from a RCT study with 34 women with POI showed a significant
decrease in systolic and diastolic blood pressure (measured in 24 hours rate) when
comparing use of HRT with the combined oral contraceptive pill at the of end of 12 months
treatment. The same trial also reported very low quality evidence on different indications of
bone density and found no significant difference in the outcomes of 25 hydroxylated Vitamin
D, 1,25 hydroxylated Vitamin D3, urinary deoxypyridinoline cross links, lumbar spine, BMD
(BMD) between women with POI who received HRT and combined oral contraceptive pill for
6 months.

The only outcomes that were found to be significantly increased with HRT compared to
combined oral contraceptive pill were osteocalcin levels and alkaline phosphate (ALP) (low
quality evidence).

Very low quality evidence from two RCTs with around 50 women with Turner Syndrome
which compared HRT with the combined oral contraceptive showed:

- no significant difference between the 2 groups in measurements of triglycerides, high
density lipoprotein (HDL) or low density (LDL) cholesterol at 6 months
- no significant difference in discontinuation rate for any cause and due to adverse events
within the 6 months of trial’s duration.

11.2.8 Evidence to recommendations

11.2.8.1 Relative value placed on the outcomes considered

The GDG considered the following outcomes to be important for their decision making: in
terms of biological markers, bone density and cardio/metabolic risk markers (insulin
resistance/lipids) were selected, changes in menopausal symptoms (including vasomotor
and sexual function), health related quality-of-life, adverse effects such as VTE comprising
DVT and PE or breast tenderness, and discontinuation for any reason.

11.2.8.2 Consideration of clinical benefits and harms

Randomised evidence showed that there was a significantly small decrease in both diastolic
and systolic blood pressure with the use of HRT for women with POI compared to combined
oral contraceptive pill. The GDG discussed that although this decrease may be small, the
long term impact on protection against CVD may be very important.

In terms of the bone density outcomes, a wide range of measurements were reported: 25
hydroxylated Vitamin D, 1,25 hydroxylated Vitamin D3, urinary deoxypyridinoline cross links,
lumbar spine BMD, osteocalcin levels and bone ALP. Only the osteocalcin levels and bone
ALP were found significantly increased with HRT compared to combined oral contraceptive
pill. However, the group discussed the results of this study and concluded that the benefit
found in the HRT group in relation to bone outcomes may not reflect a real benefit because
the women included in this study had low vitamin D at baseline which does impact on bone
health. However, given the high risk of osteoporosis for women with POI, any benefit on
bone health is of paramount importance for these women.

The only evidence that was found specifically for women with Turner Syndrome showed no
significant difference in the outcomes of triglycerides, HDL or LDL cholesterol and
discontinuation between the comparison groups.

The Group also discussed the importance of informing women with POI (including Turner
Syndrome) that HRT is not a contraceptive and therefore perimenopausal women may
require appropriate contraception.
The GDG concluded that there is very limited evidence for differences in any of the outcomes reported for the treatments of HRT and combined oral contraceptive pill, so both choices should be offered to women with POI by taking into account their preferences and needs. They recognise that the combined oral contraceptive pill is commonly taken by young women (premenopausal) and therefore it may be the preferred choice of women with POI at this age.

11.2.8.3 Consideration of economic benefits and harms

There is insufficient evidence to show whether HRT or the combined oral contraceptive pill is more effective for women with POI and both are relatively low cost interventions and therefore the GDG felt that either could be offered.

11.2.8.4 Quality of evidence

The quality of evidence from both included randomised studies was low to very low due to high risk of bias (both studies were unblinded) and imprecision. Both studies were cross-over trials and there may be residual or carry-over effect of treatments from one period to another. HRT preparations used in trials may not always represent routine clinical practice. In addition the studies were of small sample size and none of them had a longer treatment duration than a year. Therefore, the results of these studies should be interpreted with caution and the generalisation of its conclusions is under doubt.

11.2.8.5 Other considerations

The recommendations were based on both the interpretation of clinical evidence reviewed and on GDG expert opinion. The list of recommendations in this section are derived from both the clinical evidence reviewed and GDG expert opinion. No information was given in relation to women’s experience of short term symptoms. The lack of good quality clinical data makes it difficult to draw definitive conclusions whether the OCP or HRT is a better choice for women with POI.

In the absence of long term randomised prospective clinical trial data conclusions have to be drawn from clinical experience, limited short term data and observational data.

The choice to use the OCP rather than HRT is often made from the pragmatic requirement for ongoing contraception and familiarity with the pill in young women.

11.2.8.6 Key conclusions

The guideline GDG concluded that:

- there is insufficient evidence to show whether HRT or the combined oral contraceptive pill is more effective for women with POI
- there is limited evidence on the beneficial role that HRT may have on reducing systolic or diastolic blood pressure compared to OCP

11.2.9 Recommendations

59. Offer sex steroid replacement with a choice of HRT or a combined oral contraceptive to women with premature ovarian insufficiency, unless contraindicated (for example, in women with hormone-sensitive cancer).

60. Explain to women with premature ovarian insufficiency:
61. Give women with premature ovarian insufficiency and contraindications to hormonal treatments advice on bone and cardiovascular health, and symptom management (see also section 7).

### 11.2.10 Research recommendations

<table>
<thead>
<tr>
<th>Research question</th>
<th>7. What are the main clinical manifestations of premature ovarian insufficiency and the short- and long-term impact of the most common therapeutic interventions?</th>
</tr>
</thead>
</table>
| Importance to 'patients' or the population | Women with premature ovarian insufficiency can experience the effects of menopause for most of their adult life. This can lead to reduced quality of life and an increased risk of osteoporosis, cardiovascular disease and probably also dementia. There is uncertainty about the diagnosis, time course and management of premature ovarian insufficiency. For example, it is possible that different interventions produce different outcomes in terms of quality of life, and bone, cardiovascular and brain protection. Combined oral contraceptives are often prescribed when this might not be the best treatment in terms of quality of life and preservation of bone density and cardiovascular health. Short- and long-term outcomes of HRT versus combined oral contraceptives in women with premature ovarian insufficiency therefore need to be investigated. Development of a collaborative premature ovarian insufficiency registry would allow the collection of high-quality demographic, biobank (genomic) and clinical data in order to clarify:
  - the diagnosis and presentation of premature ovarian insufficiency
  - the impact of therapeutic interventions such as combined oral contraceptives, HRT and androgens
  - the long-term impact of premature ovarian insufficiency on bone density and fracture, and cardiovascular and cognitive health. |
<p>| Relevance to NICE guidance | High relevance: The NICE recommendations on HRT versus combined oral contraceptive for POI have been formulated using data from only 2 small prospective RCTs. Better quality data are urgently needed to optimise the management of young women with POI and therefore their short term quality of life and long term morbidity and mortality. |
| Relevance to the NHS | Optimised recommendations would guide NHS resource allocation and the strategic planning of care for these young women. Improved long term health in women with POI would reduce the burden on NHS resources e.g. from osteoporosis and cardiovascular related morbidity and mortality and to society in general. |
| National priorities | This was identified as a priority area by the British Menopause Society in the recommendation paper submitted to the Department of Health as part of the consultation process initiated by the Coalition Government White Paper to modernise the National Health Service. |
| Current evidence base | The current evidence base is lacking, with only 2 RCTs published thus far on the OCP versus HRT and poor quality observational and case control data. |</p>
<table>
<thead>
<tr>
<th>Research question</th>
<th>7. What are the main clinical manifestations of premature ovarian insufficiency and the short- and long-term impact of the most common therapeutic interventions?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>These data have been analysed in the guideline sections on diagnosis of POI and treatment of POI with HRT versus combined oral contraceptive.</td>
</tr>
<tr>
<td>Equality</td>
<td>Women with POI constitute a small but significant percentage of the population whose emotional and physical needs have been largely neglected by health services. Only a small proportion of units offer adequate health care professional expertise and evidence based management. The group of women with iatrogenically created POI is growing due to increasingly successful surgical, chemo and radio-therapeutic interventions; further development of survivorship programmes with due care and attention to POI should be a NHS priority.</td>
</tr>
<tr>
<td>Feasibility</td>
<td>In a sufficiently powered study, good quality data should be available within the first 5 years of a RCT of OCP versus HRT on outcomes such as quality of life, cardiovascular and osteoporosis risk markers. Longer term observational data are equally important to assess major outcomes such as CVD, fractures and cognitive functioning. It would be unethical not to offer hormonal treatment to women diagnosed with POI but a “no treatment” arm would be possible in those women wishing to avoid hormone therapy and in those in whom hormone therapy would be contraindicated. The POI registry would start yielding data immediately from amalgamation of retrospectively collated information; good quality prospectively gathered clinical and biobank data will take longer to acquire.</td>
</tr>
<tr>
<td>Other comments</td>
<td>With support from the BMS and the RCOG, a POI PICO has been submitted to the NIHR HTA for consideration but as yet this has not been made a priority area for research.</td>
</tr>
</tbody>
</table>
12 References

Transdermal HRT Investigators Group, 1993
A randomized study to compare the effectiveness, tolerability, and acceptability of two different transdermal estradiol replacement therapies. The Transdermal HRT Investigators Group, International Journal of Fertility and Menopausal Studies, 38, 5-11, 1993

Scientific Advisory Board of the Osteoporosis Society of Canada, 1996

PEPI 1996a
Effects of hormone therapy on bone mineral density: results from the postmenopausal estrogen/progestin interventions (PEPI) trial. The Writing Group for the PEPI, JAMA, 276, 1389-1396, 1996

Anonymous, 2005
Which alternative treatments work? Consumer Reports, 70, 39-43, 2005

Abdali 2010
Abdali,K., Khajehei,M., Tabatabaei,H.R., Effect of St John's wort on severity, frequency, and duration of hot flashes in premenopausal, perimenopausal and postmenopausal women: a randomized, double-blind, placebo-controlled study, Menopause, 17, 326-331, 2010

Adami 1999

Aguirre 2010
Aguirre,W., Chedraui,P., Mendoza,J., Ruilova,I., Gabapentin vs. low-dose transdermal estradiol for treating post-menopausal women with moderate to very severe hot flushes, Gynecological Endocrinology, 26, 333-337, 2010

Aitken 1974

Al-Akoum 2009
Al-Akoum,M., Maunsell,E., Verreault,R., Provencher,L., Otis,H., Dodin,S., Effects of Hypericum perforatum (St. John's wort) on hot flashes and quality of life in perimenopausal women: a randomized pilot trial, Menopause, 16, 307-314, 2009

Al-Azzawi 1997
Menopause

References

National Collaborating Centre for Women's and Children's Health

Al-Azzawi, F., Wahab, M., Habiba, M., Akkad, A., Mason, T. Continuous combined hormone replacement therapy compared with tibolone, Obstetrics and Gynecology, 93, 258-264, 1999

Al-Azzawi, F., Buckler, H.M., United Kingdom Vaginal Ring Investigator Group, Comparison of a novel vaginal ring delivering estradiol acetate versus oral estradiol for relief of vasomotor menopausal symptoms, Climacteric, 6, 118-127, 2003


Amsterdam, J.D., Yao, Y., Mao, J.J., Soeller, I., Rockwell, K., Shults, J. Randomized, double-blind, placebo-controlled trial of Cimicifuga racemosa (black cohosh) in women with anxiety disorder due to menopause, Journal of Clinical Psychopharmacology, 29, 478-483, 2009


of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial, JAMA, 291, 1701-1712, 2004

Andrist 1998


Archer 1992


Archer 2003


Archer 2012


Archer 2014


Armitage 2007


Armstrong 1996


Aslan 2007


Aso 2012


Atkinson 2004

Avis 2008

Avis 2015

Ayers 2012

Baber 1999
Baber, R.J., Templeman, C., Morton, T., Kelly, G.E., West, L., Randomized placebo-controlled trial of an isoflavone supplement and menopausal symptoms in women, Climacteric, 2, 85-92, 1999

Bacchi-Modena 1997

Bachmann 2007

Bachmann 2008

Bachmann 2009

Bachmann 2010
Bachmann, G.A., Komi, J.O., Ospemifene Study Group., Ospemifene effectively treats vulvovaginal atrophy in postmenopausal women: results from a pivotal phase 3 study, Menopause, 17, 480-486, 2010

Bagger 2004
Bagger, Y.Z., Tanko, L.B., Alexandersen, P., Hansen, H.B., Mollgaard, A., Ravn, P., Qvist, P., Kanis, J.A., Christiansen, C., Two to 3 years of hormone replacement treatment in healthy women have long-term preventive effects on bone mass and osteoporotic fractures: the PERF study, Bone, 34, 728-735, 2004

Bakken 2004
References


Bakken 2011


Baksu 2005


Baksu 2009

Baksu, B., Baksu, A., Goker, N., Citak, S., Do different delivery systems of hormone therapy have different effects on psychological symptoms in surgically menopausal women? A randomized controlled trial, Maturitas, 62, 140-145, 2009

Balk 2002


Banks 2004


Bao 2014


Barnabei 2002


Barrett-Connor 2003

Barton 2010

Becker 2009

Bener & Falah 2014

Benson 2012

Benster 2009

Beral & MillionWomen 2003

Bertelli 2002

Biglia 2009
Biglia,N., Sgandurra,P., Peano,E., Marenco,D., Moggio,G., Bounous,V., Tomasi,Cont N., Ponzoni,R., Sismondi,P., Non-hormonal treatment of hot flushes in breast cancer survivors: gabapentin vs. vitamin E, Climacteric, 12, 310-318, 2009

Bjarnason & Christiansen 2000
Menopause
References

Bjarnason, N.H., Christiansen, C., Early response in biochemical markers predicts long-term response in bone mass during hormone replacement therapy in early postmenopausal women, Bone, 26, 561-569, 2000

Blumel 2012


Bonds 2006


Bordeleau 2010


Botteman 2004


Bove 2014


Bravata 2002

Bravata, D.M., Rastegar, A., Horwitz, R.I., How do women make decisions about hormone replacement therapy?, American Journal of Medicine, 113, 22-29, 2002

Brown 2002


Brown 2006

Brown, A., Coyle, D., Chen, S., Cumming, D., Mensinkai, S., Transdermal hormone replacement therapy patches for women with postmenopausal symptoms: economic analysis of short-term use (technology report no. 61), Ottawa: Canadian coordinating office for health technology assessment, -, 2006

Brownley 2004

placebo: differences associated with years since menopause, American Journal of Obstetrics and Gynecology, 190, 1052-1058, 2004

Brunner 2010


Buckler 2003


Burger 1998


Burr 2003


Buster 2008


Butt 2008


Byrjalsen 2000


Cano 2012


Canonico 2006

Menopause

References

1 Canonico 2008

2 Canonico 2010

3 Carmignani 2010

4 Carranza-Lira & Cortes-Fuentes 2001

5 Carranza-Lira 2004

6 Casini 2006

7 Casper & Petri 1999

8 Cauley 2003

9 Chandeying & Lamlertkittikul 2007
Chandeying, V., Lamlertkittikul, S., Challenges in the conduct of Thai herbal scientific study: efficacy and safety of phytoestrogen, pueraria mirifica (Kwao Keur Kao), phase I, in the alleviation of climacteric symptoms in perimenopausal women, Journal of the Medical Association of Thailand, 90, 1274-1280, 2007

10 Chandeying & Sangthawan 2007
Chandeying, V., Sangthawan, M., Efficacy comparison of Pueraria mirifica (PM) against conjugated equine estrogen (CEE) with/without medroxyprogesterone acetate (MPA) in the treatment of climacteric symptoms in perimenopausal women: phase III study, Journal of the Medical Association of Thailand, 90, 1720-1726, 2007

Chattha 2008


Chen 2010


Cheng 2013


Cherry 2002


Cherry 2014


Cherry 2002


Chompootweep 1993

Chompootweep, S., Tankeyoon, M., Yamarat, K., Poomsuwan, P., Dusitsin, N., The menopausal age and climacteric complaints in Thai women in Bangkok, Maturitas, 17, 63-71, 1993

Chuni & Sreeramareddy 2011


Cieraad 2006

Cieraad, D., Conradt, C., Jesinger, D., Bakowski, M., Clinical study comparing the effects of sequential hormone replacement therapy with oestradiol/dydrogesterone and conjugated equine oestrogen/norgestrel on lipids and symptoms, Archives of Gynecology and Obstetrics, 274, 74-80, 2006
<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colau, J.C., Vincent, S., Marijnen, P., Allaert, F.A.</td>
<td>2012</td>
</tr>
<tr>
<td>Connelly, M.T., Ferrari, N., Hagen, N., Inui, T.S.</td>
<td>1999</td>
</tr>
<tr>
<td>Constantine, G. D., Goldstein, S. R., Archer, D. F.</td>
<td>2015</td>
</tr>
<tr>
<td>Coope, J., Thomson, J.M., Poller, L.</td>
<td>1975</td>
</tr>
<tr>
<td>Cooper, G.S., Baird, D.D.</td>
<td>1995</td>
</tr>
<tr>
<td>Coulam, C.B., Adamson SC, Annegers JF.</td>
<td>1986</td>
</tr>
<tr>
<td>Coyle, 2003</td>
<td></td>
</tr>
</tbody>
</table>
Coyle, D., Cranney, A., Tugwell, P., Economic evaluation of norethisterone acetate/ethinylestradiol (FemHRT) for women with menopausal symptoms, Pharmacoeconomics, 21, 661-9, 2003

Crandall 2002a


Crisaafulli 2004


Culhane 2003

Culhane, N.S., Estrogen plus progestin may increase incidence of dementia, Journal of Family Practice, 52, 754-755, 2003

Cunha 2010


Cushman 2004

Cushman, M., Kuller, L.H., Prentice, R., Rodabough, R.J., Psaty, B.M., Stafford, R.S., Sidney, S., Rosendaal, F.R., Women's Health Initiative Investigators., Estrogen plus progestin and risk of venous thrombosis, JAMA, 292, 1573-1580, 2004

Dalais 1998


D'Anna 2007


D'Anna 2009


Darko 2001

Darko, D.A., Dornhorst, A., Kennedy, G., Mandeno, R.C., Seed, M., Glycaemic control and plasma lipoproteins in menopausal women with T2D treated with oral and transdermal combined hormone replacement therapy, Diabetes Research and Clinical Practice, 54, 157-164, 2001
Menopause
References

Davis 2001

Davis 2006

Davis 2008

de 2000

De 2000

De 2001

deLauzon-Guillain 2009

Delmas 2000

Demetrio 2011

Dennerstein 1978

Dennerstein 1993


Derman 1995

Derman,R.J., Dawood,M.Y., Stone,S., Quality of life during sequential hormone replacement therapy -- a placebo-controlled study, International Journal of Fertility and Menopausal Studies, 40, 73-78, 1995

Deschamps 2004


deSousa-Munoz & Filizola 2009


Dessole 2004a

Dessole,Salvatore, Rubattu,Giovanni, Ambrosini,Guido, Gallo,Omar, Capobianco,Giampiero, Cherchi,Pier Luigi, Marci,Roberto, Cosmi,Erich, Efficacy of low-dose intravaginal estriol on urogenital aging in postmenopausal women, Menopause (New York, N.Y.), 11, 49-56, 2004

Devor 1992

Devor,M., Barrett-Connor,E., Renvall,M., Feigal,Jr, Ramsdell,J., Estrogen replacement therapy and the risk of venous thrombosis, American Journal of Medicine, 92, 275-282, 1992

Diaby 2007


Diem 2006

Diem,S., Grady,D., Quan,J., Vittinghoff,E., Wallace,R., Hanes,V., Ensrud,K., Effects of ultralow-dose transdermal estradiol on postmenopausal symptoms in women aged 60 to 80 years, Menopause, 13, 130-138, 2006

Doubova 2012


Edington 1980

Edington,R.F., Chagnon,J.P., Steinberg,W.M., Clonidine (Dixarit) for menopausal flushing, Canadian Medical Association Journal, 123, 23-26, 1980

Eischer 2014

Ei 2011

El, Shafie K., Al, Farsi Y., Al, Zadjali N., Al, Adawi S., Al, Busaidi Z., Al, Shafae M., Menopausal symptoms among healthy, middle-aged Omani women as assessed with the Menopause Rating Scale, Menopause, 18, 1113-1119, 2011

Elfuturi 2005

Elfuturi, A., Sherif, F., Elmahaishi, M., Chrystyn, H., Two hormone replacement therapy (HRT) regimens for middle-eastern postmenopausal women, Maturitas, 52, 52-59, 2005

Elkins 2013


Endrikat 2007

Endrikat, J., Graeser, T., Mellinger, U., Ertan, K., Holz, C., A multicenter, prospective, randomized, double-blind, placebo-controlled study to investigate the efficacy of a continuous combined hormone therapy preparation containing 1mg estradiol valerate/2mg dienogest on hot flushes in postmenopausal women, Maturitas, 58, 201-207, 2007

Engel 2011


Ensrud 2012


Eriksen & Rasmussen, 1992


Espeland 2009


Espeland 2013

References

Ettinger 1996
Ettinger, B., Friedman, G.D., Bush, T., Quesenberry, C.P., Jr., Reduced mortality associated with long-term postmenopausal estrogen therapy, Obstetrics and Gynecology, 87, 6-12, 1996

Evans 2011

Ewertz 2005

Farzaneh 2013

Faure 2002
Faure, E.D., Chantre, P., Mares, P., Effects of a standardised soy extract on hot flushes: a multicenter, double-blind, randomized, placebo-controlled study, Menopause, 9, 329-334, 2002

Ferrara 2001

Ferrari 2009

Fillenbaum 2001
Fillenbaum, G.G., Hanlon, J.T., Landerman, L.R., Schmader, K.E., Impact of estrogen use on decline in cognitive function in a representative sample of older community-resident women, American Journal of Epidemiology, 153, 137-144, 2001

Folsom 1995

Forouhari 2010

Fortin 2001

Fournier 2005


Fournier 2008

Fournier, A., Berrino, F., Clavel-Chapelon, F., Unequal risks for breast cancer associated with different hormone replacement therapies: Results from the E3N cohort study, Breast Cancer Research and Treatment, 107, 103-111, 2008

Fox-Young 1995


Freedman 2011


Freedman 2010

Freedman, R.R., Treatment of menopausal hot flashes with 5-hydroxytryptophan, Maturitas, 65, 383-385, 2010

Freeman 2011


Frisk 2008


Frisk 2012

Frisk, J., Kallstrom, A.C., Wall, N., Fredrikson, M., Hammar, M., Acupuncture improves health-related quality-of-life (HRQoL) and sleep in women with breast cancer and hot flushes, Supportive Care in Cancer, 20, 715-724, 2012

Furuhjelm 1984


Galen 2004

1 Garcia 2010
2 Garcia, J.T., Gonzaga, F., Tan, D., Ng, T.Y., Oei, P.L., Chan, C.W., Use of a multibotanical (Nutrafem) for the relief of menopausal vasomotor symptoms: a double-blind, placebo-controlled study, Menopause, 17, 303-308, 2010

Gast 2011

Gausssoin 2012

Geller 2009

Genant 1997

Giacobbe 2004
7 Giacobbe, M., Mendes Pinto-Neto, A., Simoes Costa-Paiva, L.H., Martinez, E.Z., The usefulness of ovarian volume, antral follicle count and age as predictors of menopausal status, Climacteric, 7, 255-260, 2004

Giuseppe 2007
8 Giuseppe, L., Attilio, G., Edoardo, D.N., Loredana, G., Cristina, L., Vincenzo, L., Ovarian function after cancer treatment in young women affected by Hodgkin disease (HD), Hematology, 12, 141-147, 2007

Gold 2000

Goldstein 2014

Good 1996

Good 1999


Good 1999a

Good, W.R., John, V.A., Ramirez, M., Higgins, J.E., Comparison of Alora(TM) estradiol matrix transdermal delivery system with oral conjugated equine estrogen therapy in relieving menopausal symptoms, Climacteric, 2, 29-36, 1999

Gordon 2004


Gordon 2006

Gordon, P.R., Kerwin, J.P., Boesen, K.G., Senf, J., Sertraline to treat hot flashes: a randomized controlled, double-blind, crossover trial in a general population, Menopause, 13, 568-575, 2006

Grady 2007


Graff-Iversen 2004


Griesser 2012

Griesser, H., Skonietzki, S., Fischer, T., Fielder, K., Suesskind, M., Low dose estriol pessaries for the treatment of vaginal atrophy: a double-blind placebo-controlled trial investigating the efficacy of pessaries containing 0.2mg and 0.03mg estriol, Maturitas, 71, 360-368, 2012

Grodstein 1996


Grodstein 1996a


Grodstein 1997
Menopause

References


Grodstein 2000

Grodstein, F., Manson, J.E., Stampfer, M.J., A prospective, observational study of postmenopausal hormone therapy and primary prevention of cardiovascular disease, Annals of Internal Medicine, 133, 933-941, 2000

Grodstein 2006

Grodstein, F., Manson, J.E., Stampfer, M.J., Hormone therapy and coronary heart disease: the role of time since menopause and age at hormone initiation, Journal of Women's Health, 15, 35-44, 2006

Grodstein 2008

Grodstein, F., Manson, J.E., Stampfer, M.J., Rexrode, K., Postmenopausal hormone therapy and stroke: role of time since menopause and age at initiation of hormone therapy, Archives of Internal Medicine, 168, 861-866, 2008

Guttmann 2001


Guttuso 2003


Guttuso 2003a

Guttuso, T., Jr., Kurlan, R., McDermott, M.P., Kieburtz, K., Gabapentin's effects on hot flashes in postmenopausal women: a randomized controlled trial, Obstetrics and Gynecology, 101, 337-345, 2003

Guttuso 2009

Guttuso, Jr, McDermott, M.P., Ng, P., Kieburtz, K., Effect of L-methionine on hot flashes in postmenopausal women: A randomized controlled trial, Menopause, 16, 1004-1008, 2009

Guttuso 2009a

Guttuso, T., Jr., McDermott, M.P., Ng, P., Kieburtz, K., Effect of L-methionine on hot flashes in postmenopausal women: a randomized controlled trial, Menopause, 16, 1004-1008, 2009

Haas 1988

Haas, S., Walsh, B., Evans, S., Krache, M., Ravnikar, V., Schiﬀ, I., The effect of transdermal estradiol on hormone and metabolic dynamics over a six-week period, Obstetrics and Gynecology, 71, 671-676, 1988

Hachul 2008


Haimov-Kochman, R., Barak-Glantz, E., Arbel, R., LeefsmMa, M., Brzezinski, A., Milwidsky, A., Hochner-Celnikier, D., Gradual discontinuation of hormone therapy does not prevent the reappearance of climacteric symptoms: a randomized prospective study, Menopause, 13, 370-376, 2006

Haines, C., Yu, S.L., Hiemeyer, F., Schaefers, M., Micro-dose transdermal estradiol for relief of hot flushes in postmenopausal Asian women: a randomized controlled trial, Climacteric, 12, 419-426, 2009

Hallowell, N., A qualitative study of the information needs of high-risk women undergoing prophylactic oophorectomy, Psycho-Oncology, 9, 486-495, 2000

Hammar, M., Christau, S., Nathorst-Boos, J., Rud, T., Garre, K., A double-blind, randomised trial comparing the effects of tibolone and continuous combined hormone replacement therapy in postmenopausal women with menopausal symptoms, British Journal of Obstetrics and Gynaecology, 105, 904-911, 1998


Hartley, D.E., Elsabagh, S., File, S.E., Gincosan (a combination of Ginkgo biloba and Panax ginseng): the effects on mood and cognition of 6 and 12 weeks' treatment in postmenopausal women, Nutritional Neuroscience, 7, 325-333, 2004

Haskell & Richardson 2004
Haskell, S.G., Richardson, E.D., The effect of raloxifene on cognitive function in postmenopausal women: a randomized clinical trial, Connecticut Medicine, 68, 355-358, 2004

Hassa 2010

Hassa, H., Tanir, H.M., Oge, T., Is placebo as effective as estrogen regimens on vasomotor symptoms in women with surgical menopause?, Clinical and Experimental Obstetrics and Gynecology, 37, 135-137, 2010

Hedblad 2002


Hedrick 2009

Hedrick, R.E., Ackerman, R.T., Koltun, W.D., Halvorsen, M.B., Lambrecht, L.J., Transdermal estradiol gel 0.1% for the treatment of vasomotor symptoms in postmenopausal women, Menopause, 16, 132-140, 2009

Heinrich & Wolf, 2005

Heinrich, A.B., Wolf, O.T., Investigating the effects of estradiol or estradiol/progesterone treatment on mood, depressive symptoms, menopausal symptoms and subjective sleep quality in older healthy hysterectomized women: a questionnaire study, Neuropsychobiology, 52, 17-23, 2005

Heiss 2008


Henderson 2000


Henderson 2005


Henrich 2006


Hernandez & Pluchino 2003


Hernandez 1990

Herrington 2002

Hervik & Mjaland 2009
Hervik, J., Mjaland, O., Acupuncture for the treatment of hot flashes in breast cancer patients, a randomized, controlled trial, Breast Cancer Research and Treatment, 116, 311-316, 2009

Hitchcock & Prior 2012

Ho 1999

Hogervorst 1999

Hogervorst 1999a
Hogervorst, E., Boshuisen, M., Riedel, W., Willeken, C., Jolles, J., The effect of hormone replacement therapy on cognitive function in elderly women, Psychoneuroendocrinology, 24, 43-68, 1999

Hoibraaten 2000
Hoibraaten, E., Qvigstad, E., Arnesen, H., Larsen, S., Wickstrom, E., Sandset, P.M., Increased risk of recurrent venous thromboembolism during hormone replacement therapy--results of the randomized, double-blind, placebo-controlled estrogen in venous thromboembolism trial (EVTET), Thrombosis and Haemostasis, 84, 961-967, 2000

Hoidrup 1999

Holmberg 2008

Holst 1989
Holst, J., Backstrom, T., Hammarback, S., von, Schoultz B., Progestogen addition during oestrogen replacement therapy--effects on vasomotor symptoms and mood, Maturitas, 11, 13-20, 1989

Honjo & Taketani 2009

Honjo, H., Taketani, Y., Low-dose estradiol for climacteric symptoms in Japanese women: a randomized, controlled trial, Climacteric, 12, 319-328, 2009

Honkanen 2000


Hosking 1998


Huber 2002


Hulley 2002


Hundrup 2004

Hundrup, Y. A., Hoidrup, S., Ekholm, O., Daviden, M., Obel, E. B., Risk of low-energy hip, wrist, and upper arm fractures among current and previous users of hormone replacement therapy: The Danish Nurse Cohort Study, European Journal of Epidemiology, 19, 1089-1095, 2004

Huopio 2000


Inan 2005

Inan, I., Kelekci, S., Yilmaz, B., Psychological effects of tibolone and sequential estrogen-progestogen therapy in perimenopausal women, Gynecological Endocrinology, 20, 64-67, 2005

Jackson 2006


**Jacobson 2001**


**Jadoul 2011**


**Jenks 2012**


**Jernstrom 2003**

Jernstrom, H., Bendahl, P.O., Lidfeldt, J., Nerbrand, C., Agardh, C.D., Samsioe, G., A prospective study of different types of hormone replacement therapy use and the risk of subsequent breast cancer: The women's health in the Lund area (WHILA) study (Sweden), Cancer Causes and Control, 14, 673-680, 2003

**Joffe 2010**


**Joffe 2014**


**Johnson 2004**


**Kalay 2007**

Kalay, A.E., Demir, B., Haberal, A., Kalay, M., Kandemir, O., Efficacy of citalopram on climacteric symptoms, Menopause, 14, 223-229, 2007

**Kang & Grodstein 2012**

Kang, J.H., Grodstein, F., Postmenopausal hormone therapy, timing of initiation, APOE and cognitive decline, Neurobiology of Aging, 33, 1129-1137, 2012
Kang 2004
Kang, J.H., Weuve, J., Grodstein, F., Postmenopausal hormone therapy and risk of cognitive decline in community-dwelling aging women, Neurology, 63, 101-107, 2004

Kapur 2009
Kapur, P., Sinha, B., Pereira, B.M., Measuring climacteric symptoms and age at natural menopause in an Indian population using the Greene Climacteric Scale, Menopause, 16, 378-384, 2009

Karp 2012

Katz 2007

Kawas 1997

Kenemans 2009

Kenny 2005

Kernohan 2007

Kessel 2003

Khaodhiar 2008
Khaodhiar, L., Ricciotti, H.A., Li, L., Pan, W., Schickel, M., Zhou, J., Blackburn, G.L., Daidzein-rich isoflavone aglycones are potentially effective in reducing hot flashes in menopausal women, Menopause, 15, 125-132, 2008

Khoo 1998

Khoo 2010
Khoo, S.K., O'Neil, S., Byrne, G., King, R., Travers, C., Tripcony, L., Postmenopausal hormone therapy and cognition: effects of timing and treatment type, Climacteric, 13, 259-264, 2010

Kiatponsan 2014
Kiatponsan, S., Carlson, K., Feibelmann, S., Sepucha, K., Decision aid reduces misperceptions about hormone therapy: a randomized controlled trial, Menopause, 21, 33-38, 2014

Kim 2010

Kim 2011

Kimmick 2006
Kimmick, G.G., Lovato, J., McQuellon, R., Robinson, E., Muss, H.B., Randomized, double-blind, placebo-controlled, crossover study of sertraline (Zoloft) for the treatment of hot flashes in women with early stage breast cancer taking tamoxifen, Breast Journal, 12, 114-122, 2006

Knight 1999

Knight 2001

Kokcu 2000

Komesaroff 2001
Komesaroff, P.A., Black, C.V., Cable, V., Sudhir, K., Effects of wild yam extract on menopausal symptoms, lipids and sex hormones in healthy menopausal women, Climacteric, 4, 144-150, 2001
Komulainen 1998

Kotsopoulos 2000

Kupfersztain 2003

Lacroix 2011

Lafferty & Fiske 1994
Lafferty,F.W., Fiske,M.E., Postmenopausal estrogen replacement: a long-term cohort study, American Journal of Medicine, 97, 66-77, 1994

Laliberte 2011
Laliberte,F., Dea,K., Duh,M.S., Kahler,K.H., Rolli,M., Lefebvre,P., Does the route of administration for estrogen hormone therapy impact the risk of venous thromboembolism? Estradiol transdermal system versus oral estrogen-only hormone therapy, Menopause, 18, 1052-1059, 2011

Landgren 2005
Landgren,M.B., Helmond,F.A., Engelen,S., Tibolone relieves climacteric symptoms in highly symptomatic women with at least seven hot flushes and sweats per day, Maturitas, 50, 222-230, 2005

Lando 1999
Lando,J.F., Heck,K.E., Brett,K.M., Hormone replacement therapy and breast cancer risk in a nationally representative cohort, American Journal of Preventive Medicine, 17, 176-180, 1999

Langrish 2009

Lee 2007

Lee 2010


Lees & Stevenson 2001

Lees, B., Stevenson, J.C., The prevention of osteoporosis using sequential low-dose hormone replacement therapy with estradiol-17 beta and dydrogesterone, Osteoporosis International, 12, 251-258, 2001

Legare 2007


Legare 2008


Lekander 2009


Lekander 2009a


Levis 2010

Levis, S., Strickman-Stein, N., Doerge, D.R., Krischer, J., Design and baseline characteristics of the soy phytoestrogens as replacement estrogen (SPARE) study--a clinical trial of the effects of soy isoflavones in menopausal women, Contemporary Clinical Trials, 31, 293-302, 2010

Levis 2006


Li 2006a


Liao & Hunter 1998

Lin 2011

Lindh-Astrand 2004
Lindh-Astrand, L., Nedstrand, E., Wyon, Y., Hammar, M., Vasomotor symptoms and quality of life in previously sedentary postmenopausal women randomised to physical activity or estrogen therapy, Maturitas, 48, 97-105, 2004

Lindh-Astrand 2010

Lindsay & Hart 1978
Lindsay, R., Hart, D.M., Failure of response of menopausal vasomotor symptoms to clonidine, Maturitas, 1, 21-25, 1978

Lipovac 2012

Liu & Muse 2005
Liu, J.H., Muse, K.N., The effects of progestins on bone density and bone metabolism in postmenopausal women: a randomized controlled trial, American Journal of Obstetrics and Gynecology, 192, 1316-1323, 2005

Lobo 1984

Lobo 2009

Loibl 2007

Lokkegaard 2008
References


2. Loprinzi 1994


4. Loprinzi 2000


6. Loprinzi 2002


8. Loprinzi 2006


10. Loprinzi 2010


12. Lufkin 1992


16. Luoto 2012


18. Maartens 2001

Maddalozzo 2004
Maddalozzo, G.F., Cardinal, B.J., Li, F., Snow, C.M., The association between hormone therapy use and changes in strength and body composition in early postmenopausal women, Menopause, 11, 438-446, 2004

Mahon & Williams 2000
Mahon, S.M., Williams, M., Information needs regarding menopause. Results from a survey of women receiving cancer prevention and detection services, Cancer Nursing, 23, 176-185, 2000

Manjer 2001

Mann 2012

Manson 1992

Manson 2003

Manson 2013

Manson 2013a
during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials, JAMA, 310, 1353-1368, 2013

Marsden 2001

Maxim 1995

McKenzie 2003

Meeuwsen 2002

Melton 1996

Meuwissen 2001

Middleton & Steel 2007

Mills 1989
Mills,P.K., Beeson,W.L., Phillips,R.L., Fraser,G.E., Prospective study of exogenous hormone use and breast cancer in Seventh-day Adventists, Cancer, 64, 591-597, 1989

Mingo 2000
Mingo,C., Herman,C.J., Jasperse,M., Women's stories: Ethnic variations in women's attitudes and experiences of menopause, hysterectomy, and hormone replacement therapy, Journal of Women's Health and Gender-Based Medicine, 9, S27-S38, 2000

Mirabi & Mojab 2013

Mitchell 2003
Mitchell, J.L., Cruickshanks, K.J., Klein, B.E., Palta, M., Nondahl, D.M., Postmenopausal hormone therapy and its association with cognitive impairment, Archives of Internal Medicine, 163, 2485-2490, 2003

Moriyama 2008

Morrison 2004

Mosekilde 2000

Murrkies 1995

Murrkies 2008

Murray 2001

Nachtagall 1979a

Nagamani 1987

Nahas 2007
Nathorst-Boos 2006
Nathorst-Boos, J., Floter, A., Jarkander-Rolff, M., Carlstrom, K., Schoultz, Bv. Treatment with percutaneous testosterone gel in postmenopausal women with decreased libido--effects on sexuality and psychological general well-being, Maturitas, 53, 11-18, 2006

Nedeljkovic 2014

Nedstrand 2006

Nielsen 2006

Nijland 2008

Nir 2007

Notelovitz & Mattox 2000
Notelovitz, M., Mattox, J.H., Suppression of vasomotor and vulvovaginal symptoms with continuous oral 17beta-estradiol, Menopause, 7, 310-317, 2000

Notelovitz 2000

Notelovitz 2000a
Notelovitz, M., Cassel, D., Hille, D., Furst, K.W., Dain, M.P., VandePol, C., Skarinsky, D., Efficacy of continuous sequential transdermal estradiol and norethindrone acetate in relieving vasomotor symptoms associated with menopause, American Journal of Obstetrics and Gynecology, 182, 7-12, 2000

O’Brien 2010
Odmark 2004
Odmark, I.S., Backstrom, T., Jonsson, B., Bixo, M., Well-being at onset of hormone replacement therapy: comparison between two continuous combined regimens, Climacteric, 7, 92-102, 2004

Ohira 2010

Oktem 2007

Olie 2011

Ozsoy 2002

Paganini-Hill 1991

Paganini-Hill 2005

Painovich 2012

Palacios 2004

Panay 2007

Pandya 2005
Menopause
References


8. Pinkerton, J.V., Utian, W.H., Constantine, G.D., Olivier, S., Pickar, J.H., Relief of vasomotor symptoms with the tissue-selective estrogen complex containing bazedoxifene/conjugated estrogens: a randomized, controlled trial, Menopause, 16, 1116-1124, 2009


18. Pinkerton, J.V., Utian, W.H., Constantine, G.D., Olivier, S., Pickar, J.H., Relief of vasomotor symptoms with the tissue-selective estrogen complex containing bazedoxifene/conjugated estrogens: a randomized, controlled trial, Menopause, 16, 1116-1124, 2009


Menopause

References


2. Polo-Kantola 1999


4. Pornel & Spielmann 2005

5. Pornel, B., Spielmann, D., A study of the control of climacteric symptoms in postmenopausal women following sequential regimens of 1 mg 17beta-estradiol and trimetgestone compared with a regimen containing 1 mg estradiol valerate and norethisterone over a two-year period. Gynecological Endocrinology, 21, 74-81, 2005.

6. Pornel 1995


8. Portman 2013


10. Portman 2014


12. Post 2002


14. Prentice 2009

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Punyahotra 1997</td>
<td>1997</td>
</tr>
<tr>
<td>Purdie 1995</td>
<td>1995</td>
</tr>
<tr>
<td>Qu 2009</td>
<td>2009</td>
</tr>
<tr>
<td>Randell 2002</td>
<td>2002</td>
</tr>
<tr>
<td>Rasgon 2014</td>
<td>2014</td>
</tr>
<tr>
<td>Ravn 1999</td>
<td>1999</td>
</tr>
<tr>
<td>Reddy 2006</td>
<td>2006</td>
</tr>
<tr>
<td>Reid 2004</td>
<td>2004</td>
</tr>
<tr>
<td>Ribom 2002</td>
<td>2002</td>
</tr>
<tr>
<td>Rice 2000</td>
<td>2000</td>
</tr>
<tr>
<td>Rice, M.M., Graves, A.B., McCurry, S.M., Gibbons, L.E., Bowen, J.D., McCormick, W.C., Larson, E.B., Postmenopausal estrogen and estrogen-progestin use and 2-year rate of</td>
<td></td>
</tr>
</tbody>
</table>
cognitive change in a cohort of older Japanese American women: The Kame Project, Archives of Internal Medicine, 160, 1641-1649, 2000

Roberts 1991

Ross 1999

Rossouw 2007

Rostom 2002

Rotem & Kaplan 2007
Rotem, C., Kaplan, B., Phyto-Female Complex for the relief of hot flushes, night sweats and quality of sleep: randomized, controlled, double-blind pilot study, Gynecological Endocrinology, 23, 117-122, 2007

Rothert 1997

Rovati 2000

Rozenbaum 1996

Rozenbaum 2002
Rozenbaum, H., Chevallier, O., Moyal, M., Durand, G., Perineau, M., This, P., Aerodiol study group., Efficacy and tolerability of pulsed estrogen therapy: a 12-week double-blind placebo-controlled study in highly symptomatic postmenopausal women, Climacteric, 5, 249-258, 2002

Rudolph 2004
Menopause

References

Rudolph, I., Palombo-Kinne, E., Kirsch, B., Mellinger, U., Breitbarth, H., Graser, T., Influence of a continuous combined HRT (2 mg estradiol valerate and 2 mg dienogest) on postmenopausal depression, Climacteric, 7, 301-311, 2004

Russo & Corosu 2003


Rutanen 2003


Ryan 2009


Saensak 2013


Sarrel 1998


Saxena 2010


Scambia 2000


Schairer 2000


Scharf 2007

Scharf, M.B., Berkowitz, D.V., Reape, K.Z., Effects of synthetic conjugated estrogens A on sleep quality in postmenopausal women with nocturnal diaphoresis and/or hot flushes: a pilot study, Fertility and Sterility, 88, 654-656, 2007

Schierbeck 2012

Schierbeck 2012a


Schmidt 2000


Schurmann 2004

Schurmann, R., Holler, T., Benda, N., Estradiol and drospirenone for climacteric symptoms in postmenopausal women: a double-blind, randomized, placebo-controlled study of the safety and efficacy of three dose regimens, Climacteric, 7, 189-196, 2004

Schuurman 1995


Shahnazi 2013


Shao 2012


Shin 2008


Shlipak 2001


Shumaker 2003

Shumaker 2004

Siddle 1990

Sierra 2005

Silva 2011
Silva, B.H., Martinez, D., Wender, M.C., A randomized, controlled pilot trial of hormone therapy for menopausal insomnia, Archives of Women's Mental Health, 14, 505-508, 2011

Simbalista 2010

Simon 2001

Simon 2005

Simon 2007

Simon 2008

Simon 2013


Speroff & Eisenberg 2004
Speroff, L., Eisenberg, E., Estradiol vaginal rings were an effective treatment for postmenopausal vasomotor symptoms, Evidence-Based Obstetrics and Gynecology, 6, 154-155, 2004

Speroff 1996


Speroff 2006


Speroff 2003

Speroff, L., Efficacy and tolerability of a novel estradiol vaginal ring for relief of menopausal symptoms, Obstetrics and Gynecology, 102, 823-834, 2003

Sporrong 1988


St 2001


Stahlberg 2004


Stahlberg 2005


Stampfer 1985


Stearns 2003


Stellato 1998

Stellato, R., Crawford, S.L., McKinlay, S.M., Long-cope, C., Can follicle-stimulating hormone be used to define menopausal status?, Endocrine Practice, 4, 137-141, 1998
Stevens 2000

Stevenson 2010
Stevenson, J.C., Durand, G., Kahler, E., Pertynski, T., Oral ultra-low dose continuous combined hormone replacement therapy with 0.5 mg 17-oestradiol and 2.5 mg dydrogesterone for the treatment of vasomotor symptoms: results from a double-blind, controlled study, Maturitas, 67, 227-232, 2010

Stovall 2007

Stram 2011

Studd 1995

Studd 1999a

Studd 1999

Su 2012

Suckling 2010
Suckling, Jane A., Kennedy, Ray, Lethaby, Anne, Roberts, Helen, Local oestrogen for vaginal atrophy in postmenopausal women, Cochrane Database of Systematic Reviews, -, 2010

Sutherland 2001

Swift 2005
Swift, J.A., Conway, P., Purdie, D.W., A cost-utility analysis of low-dose hormone replacement therapy in postmenopausal women with an intact uterus, Current Medical Research and Opinion, 21, 2051-61, 2005

Taaffe 2005

Tang 1996

Tarim 2002

Theroux 2010
Theroux, R., Women's decision making during the menopausal transition, Journal of the American Academy of Nurse Practitioners, 22, 612-621, 2010

Thewes 2003

TheWritingGroupforthePEPITrial 1995

Thomson & Oswald 1977
Thomson, J., Oswald, I., Effect of oestrogen on the sleep, mood, and anxiety of menopausal women, British Medical Journal, 2, 1317-1319, 1977

Tice 2003

Tice 2004
Tice, J.A., Ettinger, B., Ensrud, K., Wallace, R., Blackwell, T., Cummings, S.R., Teede, H., Isoflavones from red clover were not more effective than placebo in reducing postmenopausal hot flushes, Evidence-Based Obstetrics and Gynecology, 6, 95-96, 2004
Menopause

References

Tjonneland 2004

Tjonneland, A., Christensen, J., Thomsen, B.L., Olsen, A., Overvad, K., Ewertz, M., Mellemkjær, L., Hormone replacement therapy in relation to breast carcinoma incidence rate ratios: a prospective Danish cohort study, Cancer, 100, 2328-2337, 2004

Toh 2010


Tuppurainen 1995


Uebelhack 2006


Upmalis 2000a


Ushiroyama 2005a

Ushiroyama, T., Ikeda, A., Sakuma, K., Ueki, M., Chai-hu-gui-zhi-gan-jiang-tang regulates plasma interleukin-6 and soluble interleukin-6 receptor concentrations and improves depressed mood in climacteric women with insomnia, American Journal of Chinese Medicine, 33, 703-711, 2005

Utian 1999


Utian 2004


Utian 2004a


Utian 2005

Utian 2009


van 2009


vandeWeijer & Barentsen 2002

van de Weijer, P.H., Barentsen, R., Isoflavones from red clover (Promensil) significantly reduce menopausal hot flush symptoms compared with placebo, Maturitas, 42, 187-193, 2002

VanLeusden 1993


VanPatten 2002


Veerus 2006

Veerus, P., Hovi, S.L., Fischer, K., Rahu, M., Hakama, M., Hemminki, E., Results from the Estonian postmenopausal hormone therapy trial [ISRCTN35338757], Maturitas, 55, 162-173, 2006

Veerus 2008


Veerus 2012


Venzke 2010

Venzke, L., Calvert, J.F., Jr., Gilbertson, B., A randomized trial of acupuncture for vasomotor symptoms in post-menopausal women, Complementary Therapies in Medicine, 18, 59-66, 2010

Verhoeven 2005

Verhoeven, M.O., van der Mooren, M.J., van de Weijer, P.H., Verdegem, P.J., van der Burgt, L.M., Kenemans, P., CuraTrial Research Group., Effect of a combination of isoflavones
and Actaea racemosa Linnaeus on climacteric symptoms in healthy symptomatic perimenopausal women: a 12-week randomized, placebo-controlled, double-blind study, Menopause, 12, 412-420, 2005

**Vickers 2007**


**Vincent 2007**


**Voipio 2002**


**von & Salbach 2000**

von, Holst T., Salbach, B., Efficacy and tolerability of a new 7-day transdermal estradiol patch versus placebo in hysterectomized women with postmenopausal complaints, Maturitas, 34, 143-153, 2000

**Voss 2002**


**Walter & Britten 2002**

Walter, F.M., Britten, N., Patients’ understanding of risk: a qualitative study of decision-making about the menopause and hormone replacement therapy in general practice, Family Practice, 19, 579-586, 2002

**Walter 2004**

Walter, F.M., Emery, J.D., Rogers, M., Britten, N., Women’s views of optimal risk communication and decision making in general practice consultations about the menopause and hormone replacement therapy, Patient Education and Counseling, 53, 121-128, 2004

**Wang 2013**


**Washburn 1999**
Washburn, S., Burke, G.L., Morgan, T., Anthony, M., Effect of soy protein supplementation on serum lipoproteins, blood pressure, and menopausal symptoms in perimenopausal women, Menopause, 6, 7-13, 1999


Wimalawansa, S.J., A four-year randomized controlled trial of hormone replacement and bisphosphonate, alone or in combination, in women with postmenopausal osteoporosis, American Journal of Medicine, 104, 219-226, 1998
References

Winther 2005
Winther, K., Rein, E., Hedman, C., Femal, a herbal remedy made from pollen extracts, reduces hot flushes and improves quality of life in menopausal women: a randomized, placebo-controlled, parallel study, Climacteric, 8, 162-170, 2005

Wren & Brown 1986

Wren 2003

Wu 2001
Wu, M.H., Pan, H.A., Wang, S.T., Hsu, C.C., Chang, F.M., Huang, K.E., Quality of life and sexuality changes in postmenopausal women receiving tibolone therapy, Climacteric, 4, 314-319, 2001

Wyon 2004

Xia 2012
Xia, Y., Zhao, Y., Ren, M., Zhang, J., Wang, Y., Chang, Y., Fu, S., Fan, G., Zhu, Y., Huang, Y., Gao, X., A randomized double-blind placebo-controlled trial of a Chinese herbal medicine preparation (Jiawei Qing’e Fang) for hot flashes and quality of life in perimenopausal women, Menopause, 19, 234-244, 2012

Yang 2007

Yates 2004

Ylikangas 2007
Ylikangas, Stiina, Backstrom, Torbjorn, Heikkinen, Jorma, Cost-effectiveness of continuous combined hormone replacement therapy in long-term use: economic evaluation based on a 9-year study in Finland, Current Medical Research and OpinionCurr Med Res Opin, 23, 57-64, 2007

Yurcheshen 2009
Yurcheshen, M.E., Guttuso, T., McDermott, M., Holloway, R.G., Perlis, M., Effects of gabapentin on sleep in menopausal women with hot flashes as measured by a Pittsburgh Sleep Quality Index factor scoring model, Journal of Women’s Health, 18, 1355-1360, 2009

Zaborowska 2007

Zandi 2002

Zethraeus 1999

Zethraeus 2005
Zethraeus, Niklas, Borgstrom, Fredrik, Jonsson, Bengt, Kanis, John, Reassessment of the cost-effectiveness of hormone replacement therapy in Sweden: results based on the Women's Health Initiative randomized controlled trial, International Journal of Technology Assessment in Health Care, 21, 433-441, 2005

Zhang 2002

Zheng 2013

Zhong 2013

Ziaei 2007
Ziaei, S., Kazemnejad, A., Zareai, M., The effect of vitamin E on hot flashes in menopausal women, Gynecologic and Obstetric Investigation, 64, 204-207, 2007
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>Summary of a study, which may be published alone or as an introduction to a full scientific paper</td>
</tr>
<tr>
<td>Amenorrhoea</td>
<td>The absence of a woman's monthly period for an interval usually in excess of 6 months</td>
</tr>
<tr>
<td>Anxiety</td>
<td>A feeling of apprehension, fear, nervousness, or dread accompanied by restlessness or tension</td>
</tr>
<tr>
<td>Asymptomatic</td>
<td>Causes no symptoms</td>
</tr>
<tr>
<td>Attrition bias</td>
<td>Systematic differences between comparison groups for withdrawal or exclusion of participants from a study</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>The process used to prevent advance knowledge of group assignment in a RCT. The allocation process should be impervious to any influence by the individual making the allocation, by being administered by someone who is not responsible for recruiting participants</td>
</tr>
<tr>
<td>Applicability</td>
<td>How well the results of a study or NICE evidence review can answer a clinical question or be applied to the population being considered</td>
</tr>
<tr>
<td>Arm (of a clinical study)</td>
<td>Subsection of individuals within a study who receive one particular intervention, for example placebo arm</td>
</tr>
<tr>
<td>Association</td>
<td>Statistical relationship between 2 or more events, characteristics or other variables. The relationship may or may not be causal</td>
</tr>
<tr>
<td>Available case analysis</td>
<td>Analysis of data that is available for participants at the end of follow-up</td>
</tr>
<tr>
<td>Baseline</td>
<td>The initial set of measurements at the beginning of a study (after run-in period where applicable), with which subsequent results are compared</td>
</tr>
<tr>
<td>Before-and-after study</td>
<td>A study that investigates the effects of an intervention by measuring particular characteristics of a population both before and after taking the intervention, and assessing any change that occurs</td>
</tr>
<tr>
<td>Bias</td>
<td>Influences on a study that can make the results look better or worse than they really are. Bias can occur by chance, deliberately or as a result of systematic errors in the design and execution of a study. It can also occur at different stages in the research process, for example, during the collection, analysis, interpretation, publication or review of research data. For examples see selection bias, performance bias, information bias, confounding factor, and publication bias</td>
</tr>
<tr>
<td>Bilateral oophorectomy</td>
<td>The surgical removal of both ovaries</td>
</tr>
<tr>
<td>Biopsy</td>
<td>A minor surgical procedure during which a small tissue specimen is removed and examined microscopically for the presence of disease (often cancer)</td>
</tr>
<tr>
<td>Black cohosh</td>
<td>An herb, typically used in non-prescription supplement form</td>
</tr>
<tr>
<td>Body Mass Index (BMI)</td>
<td>A number calculated from a person's weight and height (kilograms/metres squared) that provides, for most people, a reliable indicator of body size. See also Obesity</td>
</tr>
<tr>
<td>Bone density or Bone mineral density (BMD)</td>
<td>The amount of mineralised tissue in a segment of bone. Measuring BMD is frequently used to evaluate bone strength and predict fracture risk. Results are reported as T-scores (comparison to the ideal BMD in healthy young adults) and Z-scores (comparison to other adults of the same age). See also DXA scan</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>A disease in which abnormal (malignant) cells in the breast divide and multiply in an uncontrolled fashion. The cells can invade nearby tissue and can spread through the bloodstream and lymphatic system (lymph nodes) to other parts of the body</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>An umbrella term used to describe many conditions related to the cardio-vascular system</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>(CVD)</td>
<td>circulatory system, both inside and outside the heart. Includes heart disease, coronary artery disease (CAD), and coronary heart disease (CHD) as well as peripheral vascular disease. See also Coronary artery disease, Heart disease and venous thrombo-embolic disease</td>
</tr>
<tr>
<td>Carer (caregiver)</td>
<td>Someone who looks after family, partners or friends in need of help because they are ill, frail or have a disability</td>
</tr>
<tr>
<td>Case–control study</td>
<td>A study to find out the cause(s) of a disease or condition. This is done by comparing a group of patients who have the disease or condition (cases) with a group of people who do not have it (controls) but who are otherwise as similar as possible (in characteristics thought to be unrelated to the causes of the disease or condition). This means the researcher can look for aspects of their lives that differ to see if they may cause the condition. Such studies are retrospective because they look back in time from the outcome to the possible causes of a disease or condition</td>
</tr>
<tr>
<td>Case series</td>
<td>Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients</td>
</tr>
<tr>
<td>Clinical audit</td>
<td>A systematic process for setting and monitoring standards of clinical care. Whereas ‘guidelines’ define what the best clinical practice should be, ‘audit’ investigates whether best practice is being carried out. Clinical audit can be described as a cycle or spiral. Within the cycle there are stages that follow a systematic process of establishing best practice, measuring care against specific criteria, taking action to improve care and monitoring to sustain improvement. The spiral suggests that as the process continues, each cycle aspires to a higher level of quality</td>
</tr>
<tr>
<td>Clinical efficacy</td>
<td>The extent to which an intervention is active when studied under controlled research conditions</td>
</tr>
<tr>
<td>Clinical effectiveness</td>
<td>How well a specific test or treatment works when used in the ‘real world’ (for example, when used by a doctor with a patient at home), rather than in a carefully controlled clinical trial. Trials that assess clinical effectiveness are sometimes called management trials. Clinical effectiveness is not the same as efficacy</td>
</tr>
<tr>
<td>Clinician</td>
<td>A healthcare professional who provides patient care. For example, a doctor, nurse or physiotherapist</td>
</tr>
<tr>
<td>Cochrane Review</td>
<td>The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of RCTs prepared by the Cochrane Collaboration)</td>
</tr>
<tr>
<td>Cognitive function</td>
<td>Conscious intellectual activity (thinking, reasoning, remembering)</td>
</tr>
<tr>
<td>Cohort study</td>
<td>A study with 2 or more groups of people – cohorts – with similar characteristics. One group receives a treatment, is exposed to a risk factor or has a particular symptom and the other group does not. The study follows their progress over time and records what happens</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>A disease or condition that someone has in addition to the health problem being studied or treated</td>
</tr>
<tr>
<td>Comparability</td>
<td>The process used to ensure that the person deciding to enter a participant into a RCT does not know the comparison group into which that individual will be allocated. This is distinct from blinding, and is aimed at preventing selection bias. Some attempts at concealing allocation are more prone to manipulation than others, and the method of allocation concealment is used as an assessment of the quality of a trial</td>
</tr>
<tr>
<td>Compounded bioidentical hormones</td>
<td>Unregulated plant-derived hormonal combinations similar or identical to human hormones that are compounded by pharmacies to the specification of the prescriber. These are not quality controlled as prescribed medication</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Confidence interval (CI)</strong></td>
<td>There is always some uncertainty in research. This is because a small group of patients is studied to predict the effects of a treatment on the wider population. The confidence interval is a way of expressing how certain we are about the findings from a study, using statistics. It gives a range of results that is likely to include the 'true' value for the population. The CI is usually stated as '95% CI', which means that the range of values has a 95 in a 100 chance of including the 'true' value. For example, a study may state that 'based on our sample findings, we are 95% certain that the 'true' population blood pressure is not higher than 150 and not lower than 110'. In such a case the 95% CI would be 110 to 150. A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment - often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example, if a large number of patients have been studied).</td>
</tr>
<tr>
<td><strong>Confounding factor</strong></td>
<td>Something that influences a study and can result in misleading findings if it is not understood or appropriately dealt with. For example, a study of heart disease may look at a group of people that exercises regularly and a group that does not exercise. If the ages of the people in the 2 groups are different, then any difference in heart disease rates between the 2 groups could be because of age rather than exercise. Therefore age is a confounding factor.</td>
</tr>
<tr>
<td><strong>Continuous outcome</strong></td>
<td>Data with a potentially infinite number of possible values within a given range. Height, weight and blood pressure are examples of continuous variables.</td>
</tr>
<tr>
<td><strong>Contraception</strong></td>
<td>Any method used to prevent pregnancy during sexual activity. Perimenopausal women who wish to avoid pregnancy are advised to use reliable contraception until 2 years have passed without a menstrual period if aged under 50, until 1 year if aged 50 or older, or until the age of 55 years (NICE publication cks.nice.org.uk/contraception-assessment)</td>
</tr>
<tr>
<td><strong>Control group</strong></td>
<td>A group of people in a study who do not receive the treatment or test being studied. Instead, they may receive the standard treatment (sometimes called 'usual care') or a dummy treatment (placebo). The results for the control group are compared with those for a group receiving the treatment being tested. The aim is to check for any differences. Ideally, the people in the control group should be as similar as possible to those in the treatment group, to make it as easy as possible to detect any effects due to the treatment</td>
</tr>
<tr>
<td><strong>Coronary artery disease (CAD)</strong></td>
<td>Sometimes called coronary heart disease (CHD). The most common form of heart disease, CAD refers to damaged or diseased blood vessels (coronary arteries) that supply blood to the heart muscle. See also Cardiovascular disease, Heart disease</td>
</tr>
<tr>
<td><strong>Cost–benefit analysis (CBA)</strong></td>
<td>Cost-benefit analysis is one of the tools used to carry out an economic evaluation. The costs and benefits are measured using the same monetary units (for example, pounds sterling) to see whether the benefits exceed the costs</td>
</tr>
<tr>
<td><strong>Cost–consequences analysis (CCA)</strong></td>
<td>Cost-consequence analysis is one of the tools used to carry out an economic evaluation. This compares the costs (such as treatment and hospital care) and the consequences (such as health outcomes) of a test or treatment with a suitable alternative. Unlike cost-benefit analysis or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (such as the quality-adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out</td>
</tr>
<tr>
<td><strong>Cost-effectiveness analysis</strong></td>
<td>Cost-effectiveness analysis is one of the tools used to carry out an economic evaluation. The benefits are expressed in non-monetary terms</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>(CEA)</td>
<td>related to health, such as symptom-free days, heart attacks avoided, deaths avoided or life years gained (that is, the number of years by which life is extended as a result of the intervention)</td>
</tr>
<tr>
<td>Cost-effectiveness model</td>
<td>An explicit mathematical framework, which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes</td>
</tr>
<tr>
<td>Cost–utility analysis (CUA)</td>
<td>Cost-utility analysis is one of the tools used to carry out an economic evaluation. The benefits are assessed in terms of both quality and duration of life, and expressed as quality-adjusted life years (QALYs). See also utility</td>
</tr>
<tr>
<td>COX proportional hazard model</td>
<td>In survival analysis, a statistical model that asserts that the effect of the study factors (for example the intervention of interest) on the hazard rate (the risk of occurrence of an event) in the study population is multiplicative and does not change over time</td>
</tr>
<tr>
<td>Credible interval (CrI)</td>
<td>The Bayesian equivalent of a confidence interval</td>
</tr>
<tr>
<td>Decision analysis</td>
<td>An explicit quantitative approach to decision-making under uncertainty, based on evidence from research. This evidence is translated into probabilities, and then into diagrams or decision trees which direct the clinician through a succession of possible scenarios, actions and outcomes</td>
</tr>
<tr>
<td>Depression</td>
<td>Altered mood characterised by severe despondency or despair, often with feelings of inadequacy or guilt, which is persistent and interferes with activities of daily living</td>
</tr>
<tr>
<td>Diabetes</td>
<td>A group of diseases in which the body cannot properly control the amount of sugar in the blood, resulting in high sugar levels that may cause a variety of complications ranging from cardiovascular disease to blindness and kidney failure. Diabetes occurs when the body does not produce enough insulin or does not use it properly (insulin resistance)</td>
</tr>
<tr>
<td>Dichotomous outcomes</td>
<td>Outcome that can take one of 2 possible values, such as dead/alive, smoker/non-smoker, present/not present (also called binary data)</td>
</tr>
<tr>
<td>Discounting</td>
<td>Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present</td>
</tr>
<tr>
<td>Dominance</td>
<td>A health economics term. When comparing tests or treatments, an option that is both less effective and costs more is said to be 'dominated' by the alternative</td>
</tr>
<tr>
<td>Drop-out</td>
<td>A participant who withdraws from a trial before the end</td>
</tr>
<tr>
<td>Economic evaluation</td>
<td>An economic evaluation is used to assess the cost effectiveness of healthcare interventions (that is, to compare the costs and benefits of a healthcare intervention to assess whether it is worth doing). The aim of an economic evaluation is to maximise the level of benefits - health effects - relative to the resources available. It should be used to inform and support the decision-making process; it is not supposed to replace the judgement of healthcare professionals. There are several types of economic evaluation: cost-benefit analysis, cost-consequence analysis, cost-effectiveness analysis, cost-minimisation analysis and cost-utility analysis. They use similar methods to define and evaluate costs, but differ in the way they estimate the benefits of a particular drug, programme or intervention</td>
</tr>
<tr>
<td>Effect (as in effect measure, treatment effect, estimate of)</td>
<td>A measure that shows the magnitude of the outcome in one group compared with that in a control group. For example, if the absolute risk reduction is shown to be 5% and it is the outcome of interest, the effect size is 5%.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>effect, effect size</td>
<td>The effect size is usually tested, using statistics, to find out how likely it is that the effect is a result of the treatment and has not just happened by chance</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>How beneficial a test or treatment is under usual or everyday conditions</td>
</tr>
<tr>
<td>Efficacy</td>
<td>How beneficial a test, treatment or public health intervention is under ideal conditions (for example, in a laboratory)</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td>Cancer of the inner lining (endometrium) of the uterus</td>
</tr>
<tr>
<td>Epidemiological study</td>
<td>The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (for example, infection, diet) and interventions</td>
</tr>
<tr>
<td>EQ-5D (EuroQol 5 dimensions)</td>
<td>A standardised instrument used to measure health-related quality-of-life. It provides a single index value for health status</td>
</tr>
<tr>
<td>Equivalence study</td>
<td>A trial designed to determine whether the response to 2 or more treatments differs by an amount that is clinically unimportant. This is usually demonstrated by showing that the true treatment difference is likely to lie between a lower and an upper equivalence level of clinically acceptable differences</td>
</tr>
<tr>
<td>Estradiol</td>
<td>Also called 17beta-estradiol. The most potent of the naturally occurring oestrogens and the primary oestrogen produced by women in their reproductive years. Available in oral, skin patch, and vaginal prescription drugs</td>
</tr>
<tr>
<td>Evidence</td>
<td>Information on which a decision or guidance is based. Evidence is obtained from a range of sources including RCTs, observational studies, expert opinion (of clinical professionals or patients)</td>
</tr>
<tr>
<td>Exclusion criteria (literature review)</td>
<td>Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence</td>
</tr>
<tr>
<td>Exclusion criteria (clinical study)</td>
<td>Criteria that define who is not eligible to participate in a clinical study</td>
</tr>
<tr>
<td>Extended dominance</td>
<td>If Option A is both more clinically effective than Option B and has a lower cost per unit of effect, when both are compared with a do-nothing alternative then Option A is said to have extended dominance over Option B. Option A is therefore more cost effective and should be preferred, other things remaining equal</td>
</tr>
<tr>
<td>Extrapolation</td>
<td>An assumption that the results of studies of a specific population will also hold true for another population with similar characteristics</td>
</tr>
<tr>
<td>False negative</td>
<td>A diagnostic test result that incorrectly indicates that an individual does not have the disease of interest, when they do actually have it</td>
</tr>
<tr>
<td>False positive</td>
<td>A diagnostic test result that incorrectly indicates that an individual has the disease of interest, when they actually do not have it</td>
</tr>
<tr>
<td>Fertile</td>
<td>Capable of reproducing</td>
</tr>
<tr>
<td>Fixed-effect model</td>
<td>In meta-analysis, a model that calculates a pooled effect estimate using the assumption that all observed variation between studies is caused by random sample variability. Studies are assumed to estimating the same overall effect</td>
</tr>
<tr>
<td>Follicle-stimulating hormone (FSH)</td>
<td>A hormone produced by the pituitary gland (located at the base of the brain). In women, FSH stimulates the growth of ovarian follicles (the small cysts that hold the eggs) and the supporting cells responsible for the growth and nurturing of the egg. FSH also stimulates production of oestrogen by the ovaries. When oestrogen production is low (e.g. after menopause), FSH levels will be high</td>
</tr>
<tr>
<td>Follow-up</td>
<td>Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related aspects</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Forest plot</td>
<td>A graphical representation of the individual results of each study included in a meta-analysis together with the combined meta-analysis result. The plot also allows readers to see the heterogeneity among the results of the studies. The results of individual studies are shown as squares centred on each study’s point estimate. A horizontal line runs through each square to show each study’s confidence interval. The overall estimate from the meta-analysis and its confidence interval are shown at the bottom, represented as a diamond. The centre of the diamond represents the pooled point estimate, and its horizontal tips represent the confidence interval.</td>
</tr>
<tr>
<td>Fracture</td>
<td>The breaking of bone, resulting either from trauma (such as a fall) or because bone has become weakened from a condition such as osteoporosis. See also Osteoporosis.</td>
</tr>
<tr>
<td>Fragility fracture</td>
<td>Fractures that result from mechanical forces that would not ordinarily result in fracture (such as a fall from a standing height or less). Reduced bone density is a major risk factor for fragility fractures, which occur most commonly in the spine, hip and wrist.</td>
</tr>
<tr>
<td>Generalisability</td>
<td>The extent to which the results of a study hold true for groups that did not participate in the research. See also external validity.</td>
</tr>
<tr>
<td>Gold standard</td>
<td>A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease.</td>
</tr>
<tr>
<td>GRADE, GRADE profile</td>
<td>A system developed by the GRADE Working Group to address the Short-comings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.</td>
</tr>
<tr>
<td>Harms</td>
<td>Adverse effects of an intervention.</td>
</tr>
<tr>
<td>Hazard ratio</td>
<td>A graphical representation of the individual results of each study included in a meta-analysis together with the combined meta-analysis result. The plot also allows readers to see the heterogeneity among the results of the studies. The results of individual studies are shown as squares centred on each study’s point estimate. A horizontal line runs through each square to show each study’s confidence interval. The overall estimate from the meta-analysis and its confidence interval are shown at the bottom, represented as a diamond. The centre of the diamond represents the pooled point estimate, and its horizontal tips represent the confidence interval.</td>
</tr>
<tr>
<td>Health economics</td>
<td>Study or analysis of the cost of using and distributing healthcare resources.</td>
</tr>
<tr>
<td>Health-related quality-of-life (HRQoL)</td>
<td>A measure of the effects of an illness to see how it affects someone's day-to-day life.</td>
</tr>
<tr>
<td>Heterogeneity</td>
<td>The term is used in meta-analyses and systematic reviews to describe when the results of a test or treatment (or estimates of its effect) differ.</td>
</tr>
<tr>
<td>Hormone replacement therapy (HRT)</td>
<td>Prescription drugs used most often when treating menopause symptoms. Encompasses both oestrogen therapy and oestrogen plus progestogen therapy.</td>
</tr>
<tr>
<td>Hot flush</td>
<td>The most common menopause-related symptom, comprising a sudden feeling of heat, resulting in a red, flushed face and neck, perspiration and a rapid heartbeat, lasting a short time and often followed by a cold chill. See also Vasomotor symptoms.</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Abnormally high blood pressure.</td>
</tr>
<tr>
<td>Hysterectomy</td>
<td>Surgical removal of the uterus. This does not necessarily involve removal of the ovaries (see Bilateral Oophorectomy).</td>
</tr>
<tr>
<td>Iatrogenic</td>
<td>Adverse consequence of medical examination treatment or advice, for</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Imprecision</td>
<td>Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect</td>
</tr>
<tr>
<td>Incidence</td>
<td>The incidence of a disease is the rate at which new cases occur in a population during a specified period</td>
</tr>
<tr>
<td>Inclusion criteria (literature review)</td>
<td>Explicit criteria used to decide which studies should be considered as potential sources of evidence</td>
</tr>
<tr>
<td>Inclusion criteria (clinical study)</td>
<td>Specific criteria that define who is eligible to participate in a clinical study</td>
</tr>
<tr>
<td>Incremental cost</td>
<td>The extra cost linked to using one test or treatment rather than another. Or the additional cost of doing a test or providing a treatment more frequently</td>
</tr>
<tr>
<td>Incremental cost effectiveness ratio (ICER)</td>
<td>The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest for one treatment compared with another</td>
</tr>
<tr>
<td>Incremental net benefit (INB)</td>
<td>The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000 x QALYs gained) – Incremental cost</td>
</tr>
<tr>
<td>Indirectness</td>
<td>The available evidence is different to the review question being addressed, in terms of PICO (population, intervention, comparison and outcome)</td>
</tr>
<tr>
<td>Induced menopause.</td>
<td>Menopause brought on by treatment, for example, surgical removal of the ovaries</td>
</tr>
<tr>
<td>Intention-to-treat analysis (ITT)</td>
<td>An assessment of the people taking part in a clinical trial, based on the group they were initially (and randomly) allocated to. This is regardless of whether or not they dropped out, fully complied with the treatment or switched to an alternative treatment. Intention-to-treat analyses are often used to assess clinical effectiveness because they mirror actual practice: that is, not everyone complies with treatment and the treatment people receive may be changed according to how they respond to it</td>
</tr>
<tr>
<td>Intervention</td>
<td>In medical terms this could be a drug treatment, surgical procedure, diagnostic or psychological therapy. Examples of public health interventions could include action to help someone to be physically active or to eat a more healthy diet</td>
</tr>
<tr>
<td>Isoflavones</td>
<td>Naturally occurring oestrogen-like compounds found in soybeans, soy products, and red clover</td>
</tr>
<tr>
<td>Kappa statistic</td>
<td>A statistical measure of inter-rater agreement that takes into account the agreement occurring by chance</td>
</tr>
<tr>
<td>Length of stay</td>
<td>The total number of days a patient stays in hospital</td>
</tr>
<tr>
<td>Licence</td>
<td>See ‘Product licence’</td>
</tr>
<tr>
<td>Life years gained</td>
<td>Mean average years of life gained per person as a result of the intervention compared with an alternative intervention</td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>The likelihood ratio combines information about the sensitivity and specificity. It tells you how much a positive or negative result changes the likelihood that a patient would have the disease. The likelihood ratio of a positive test result (LRplus) is sensitivity divided by (1 minus specificity).</td>
</tr>
<tr>
<td>Loss to follow-up</td>
<td>Patients who have withdrawn from the clinical trial at the point of follow-up</td>
</tr>
<tr>
<td>Low mood</td>
<td>Mild depression symptoms that impair quality of life but are usually intermittent and often associated with hormonal fluctuations, for example</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Markov model</td>
<td>A method for estimating long-term costs and effects for recurrent or chronic conditions, based on health states and the probability of transition between them within a given time period (cycle)</td>
</tr>
<tr>
<td>Mammogram</td>
<td>Specialised x-rays of the breast used to detect abnormal growths or changes in the breast tissue</td>
</tr>
<tr>
<td>Mean</td>
<td>An average value, calculated by adding all the observations and dividing by the number of observations</td>
</tr>
<tr>
<td>Mean difference</td>
<td>In meta-analysis, a method used to combine measures on continuous scales (such as weight), where the mean, standard deviation and sample size in each group are known. The weight given to the difference in means from each study (e.g. how much influence each study has on the overall results of the meta-analysis) is determined by the precision of its estimate of effect</td>
</tr>
<tr>
<td>Median</td>
<td>The value of the observation that comes half way when the observations are ranked in order</td>
</tr>
<tr>
<td>Menopause</td>
<td>A biological stage in a woman's life that occurs when she stops menstruating and reaches the end of her natural reproductive life. This occurs when the ovaries stop functioning, and includes the cessation of egg (oocyte) maturation and of oestrogen and progesterone secretion</td>
</tr>
<tr>
<td>Menstrual cycle</td>
<td>The cycle of changes in the uterus and ovaries during a woman’s reproductive life, resulting in menstruation, typically every 4 weeks. During the cycle an egg develops in the ovary and is released, the lining of the uterus thickens to prepare for implantation of a fertilised egg, and if this does not occur, the lining of the uterus is shed through menstruation and the cycle begins again. This cycle typically becomes irregular during perimenopause and ends completely at menopause</td>
</tr>
<tr>
<td>Meta-analysis</td>
<td>A method often used in systematic reviews. Results from several studies of the same test or treatment are combined to estimate the overall effect of the treatment</td>
</tr>
<tr>
<td>Minimal important difference (MID)</td>
<td>Threshold for clinical importance which represents the minimal important difference for benefit or for harm; e.g. the threshold at which drug A is less effective than drug B by an amount that is clinically important to patients</td>
</tr>
<tr>
<td>Monte carlo</td>
<td>A technique used to approximate the probability of certain outcomes by running multiple simulations using random variables</td>
</tr>
<tr>
<td>Multivariate model</td>
<td>A statistical model for analysis of the relationship between 2 or more predictors, (independent) variables and the outcome (dependent) variable</td>
</tr>
<tr>
<td>Net monetary benefit (NMB)</td>
<td>The value (usually in monetary terms) of an intervention net of its cost. The NMB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the NMB is calculated as: (£20,000 x QALYs gained) – cost</td>
</tr>
<tr>
<td>Network meta-analysis</td>
<td>Meta-analysis in which multiple treatments (that is, 3 or more) are being compared using both direct comparisons of interventions within RCTs and indirect comparisons across trials based on a common comparator</td>
</tr>
<tr>
<td>Night sweats</td>
<td>Hot flushes that occur at night causing heavy perspiration, often interfering with sleep</td>
</tr>
<tr>
<td>Non-inferiority trial</td>
<td>A trial designed to determine whether the effect of a new treatment is not worse than a standard treatment by more than a pre-specified amount. A one-sided version of an equivalence trial</td>
</tr>
<tr>
<td>Number needed to treat (NNT)</td>
<td>The average number of patients who need to be treated to get a positive outcome. For example, if the NNT is 4, then 4 patients would have to be treated to ensure 1 of them gets better. The closer the NNT is to 1, the better the treatment. For example, if you give a stroke prevention drug to 20 people before 1 stroke is prevented, the number needed to treat is 20.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>Obese</td>
<td>Excessive accumulation of fat in the body. Obesity is defined as a body mass index over 30 kg/m2 (WHO) and is associated with health problems including T2D, cardiovascular disease, stroke, hypertension, some cancers, and premature death. See also Body Mass Index</td>
</tr>
<tr>
<td>Observational study</td>
<td>Individuals or groups are observed or certain factors are measured. No attempt is made to affect the outcome. For example, an observational study of a disease or treatment would allow 'nature' or usual medical care to take its course. Changes or differences in one characteristic (for example, whether or not people received a specific treatment or intervention) are studied without intervening. There is a greater risk of selection bias than in experimental studies</td>
</tr>
<tr>
<td>Odds ratio (OR)</td>
<td>Odds are a way to represent how likely it is that something will happen (the probability). An odds ratio compares the probability of something in one group with the probability of the same thing in another. An odds ratio of 1 between 2 groups would show that the probability of the event (for example a person developing a disease, or a treatment working) is the same for both. An odds ratio greater than 1 means the event is more likely in the first group. An odds ratio less than 1 means that the event is less likely in the first group. Sometimes probability can be compared across more than 2 groups - in this case, one of the groups is chosen as the 'reference category', and the odds ratio is calculated for each group compared with the reference category. For example, to compare the risk of dying from lung cancer for non-smokers, occasional smokers and regular smokers, non-smokers could be used as the reference category. Odds ratios would be worked out for occasional smokers compared with non-smokers and for regular smokers compared with non-smokers. See also confidence interval, relative risk, risk ratio</td>
</tr>
<tr>
<td>Oestrogens</td>
<td>Hormonal compounds produced by the ovaries, which influence the growth and health of female reproductive organs and are active in many body tissues. The 3 main naturally occurring oestrogens in women are oestradiol (premenopausal women), oestrone, and estriol. Oestradiol levels fall after menopause. Several types of oestrogen therapies are available for treatment of menopause, and also in the combined oral contraceptive, but at higher doses than those used for menopause treatment</td>
</tr>
<tr>
<td>Oestrogen plus progestogen therapy</td>
<td>Also known as combination hormone therapy. Oestrogen is the hormone in this duo that provides the most relief for menopause-related symptoms. Progestogen is added to protect the lining of the uterus from oestrogen stimulation which increases risk of endometrial cancer if given alone</td>
</tr>
<tr>
<td>Opportunity cost</td>
<td>The loss of other health care programmes displaced by investment in or introduction of another intervention. This may be best measured by the health benefits that could have been achieved had the money been spent on the next best alternative healthcare intervention</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>A condition in which the bone density of the skeleton has decreased to a point where bone has become fragile and at higher risk for fractures, often with little or no trauma. Common amongst older women, because bone mineral loss usually occurs after menopause, which is related to the decline in estrogen levels</td>
</tr>
<tr>
<td>Outcome</td>
<td>The impact that a test, treatment, policy, programme or other intervention has on a person, group or population. Outcomes from interventions to improve the public's health could include changes in knowledge and behaviour related to health, societal changes (for example, a reduction in crime rates) and a change in people's health and wellbeing or health status. In clinical terms, outcomes could include the number of patients who fully recover from an illness or the number of hospital admissions, and an improvement or deterioration in someone's health, functional</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ability, symptoms or situation</td>
<td>Researchers should decide what outcomes to measure before a study begins</td>
</tr>
<tr>
<td>P value</td>
<td>The p value is a statistical measure that indicates whether or not an effect is statistically significant. For example, if a study comparing 2 treatments found that one seems more effective than the other, the p value is the probability of obtaining these results by chance. By convention, if the p value is below 0.05 (that is, there is less than a 5% probability that the results occurred by chance) it is considered that there probably is a real difference between treatments. If the p value is 0.001 or less (less than a 1% probability that the results occurred by chance), the result is seen as highly significant. If the p value shows that there is likely to be a difference between treatments, the confidence interval describes how big the difference in effect might be</td>
</tr>
<tr>
<td>Performance bias</td>
<td>Systematic differences between intervention groups in care provided apart from the intervention being evaluated. Blinding of study participants (both the recipients and providers of care) is used to protect against performance bias</td>
</tr>
<tr>
<td>Perimenopause</td>
<td>The time in which a woman has irregular cycles of ovulation and menstruation before the menopause up until 1 year after her final period (also known as menopausal transition or climacteric)</td>
</tr>
<tr>
<td>Phytoestrogens.</td>
<td>Plant compounds (such as isoflavones) that have a chemical structure similar to that of oestrogen and have weak oestrogen-like biologic activity. Available in foods (such as soy) and as non-prescription supplements. See also Isoflavones, Red clover</td>
</tr>
<tr>
<td>Placebo</td>
<td>A fake (or dummy) treatment given to participants in the control group of a clinical trial. It is indistinguishable from the actual treatment (which is given to participants in the experimental group). The aim is to determine what effect the experimental treatment has had - over and above any placebo effect caused because someone has received (or thinks they have received) care or attention</td>
</tr>
<tr>
<td>Placebo effect</td>
<td>A beneficial (or adverse) effect produced by a placebo and not due to any property of the placebo itself</td>
</tr>
<tr>
<td>Post-hoc analysis</td>
<td>Statistical analyses that are not specified in the trial protocol, and are generally suggested by the data</td>
</tr>
<tr>
<td>Postmenopause</td>
<td>The time after a woman has not had a period for 12 consecutive months</td>
</tr>
<tr>
<td>Power (statistical)</td>
<td>The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed</td>
</tr>
<tr>
<td>Premature ovarian insufficiency</td>
<td>Menopause occurring before the age of 40 years (also known as premature ovarian failure or premature menopause). It can occur naturally or iatrogenically (that is, as a result of treatment)</td>
</tr>
<tr>
<td>Premenopause</td>
<td>The span of time from puberty (onset of menstrual periods) to perimenopause</td>
</tr>
<tr>
<td>Prevalence</td>
<td>The prevalence of a disease is the proportion of a population that are cases at a point in time</td>
</tr>
<tr>
<td>Primary care</td>
<td>Healthcare delivered outside hospitals. It includes a range of services provided by GPs, nurses, health visitors, midwives and other healthcare professionals and allied health professionals such as dentists, pharmacists and opticians</td>
</tr>
<tr>
<td>Primary outcome</td>
<td>The outcome of greatest importance, usually the one in a study that the power calculation is based on</td>
</tr>
<tr>
<td>Product licence</td>
<td>An authorisation from the MHRA to market a medicinal product</td>
</tr>
<tr>
<td>Prognosis</td>
<td>A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Progesterone</td>
<td>A naturally occurring hormone produced by the ovaries, which acts on the lining of the uterus</td>
</tr>
<tr>
<td>Progestogen</td>
<td>A synthetic hormone virtually identical to progesterone, with similar biological effects. Several different progestogens exist and are used in hormone replacement therapy</td>
</tr>
<tr>
<td>Prospective study</td>
<td>A research study in which the health or other characteristic of participants is monitored (or 'followed up') for a period of time, with events recorded as they happen. This contrasts with retrospective studies</td>
</tr>
<tr>
<td>Protocol (review)</td>
<td>A document written prior to commencing a review that details exactly how evidence to answer a review question will be obtained and synthesised. It defines in detail the population of interest, the interventions, the comparators/controls, and the outcomes of interest (PICO)</td>
</tr>
<tr>
<td>Publication bias</td>
<td>Publication bias occurs when researchers publish the results of studies showing that a treatment works well and don't publish those showing it did not have any effect. If this happens, analysis of the published results will not give an accurate idea of how well the treatment works. This type of bias can be assessed by a funnel plot</td>
</tr>
<tr>
<td>Quality-of-life</td>
<td>See ‘Health-related quality-of-life’</td>
</tr>
<tr>
<td>Quality-adjusted life year (QALY)</td>
<td>A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality-of-life. One QALY is equal to 1 year of life in perfect health. QALYS are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality-of-life score (on a scale of 0 to 1). It is often measured in terms of the person's ability to perform the activities of daily life, freedom from pain and mental disturbance</td>
</tr>
<tr>
<td>Random effect model</td>
<td>In meta-analysis, a model that calculates a pooled effect estimate using the assumption that each study is estimating a different true treatment effect due to real differences between studies. Observed variation in effects are therefore caused by a combination of random sample variability (within-study variation) and heterogeneity between studies (between-study variation). The overall effects is an average of the estimated true study effects</td>
</tr>
<tr>
<td>Randomisation</td>
<td>Assigning participants in a research study to different groups without taking any similarities or differences between them into account. For example, it could involve using a random numbers table or a computer-generated random sequence. It means that each individual (or each group in the case of cluster randomisation) has the same chance of receiving each intervention</td>
</tr>
<tr>
<td>Randomised controlled trial (RCT)</td>
<td>A study in which a number of similar people are randomly assigned to 2 (or more) groups to test a specific drug or treatment. One group (the experimental group) receives the treatment being tested, the other (the comparison or control group) receives an alternative treatment, a dummy treatment (placebo) or no treatment at all. The groups are followed up to see how effective the experimental treatment was. Outcomes are measured at specific times and any difference in response between the groups is assessed statistically. This method is also used to reduce bias</td>
</tr>
<tr>
<td>Red clover</td>
<td>A member of the legume plant family rich in phytoestrogens</td>
</tr>
<tr>
<td>Reference standard</td>
<td>The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice</td>
</tr>
<tr>
<td>Relative risk (RR)</td>
<td>The ratio of the risk of disease or death among those exposed to certain conditions compared with the risk for those who are not exposed to the same conditions (for example, the risk of people who smoke getting lung cancer compared with the risk for people who do not smoke).</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-----------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>Reporting bias</td>
<td>See ‘Publication bias’</td>
</tr>
<tr>
<td>Resource implication</td>
<td>The likely impact in terms of finance, workforce or other NHS resources.</td>
</tr>
<tr>
<td>Retrospective study</td>
<td>A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected</td>
</tr>
<tr>
<td>Review question</td>
<td>The plan or set of steps to be followed in a study. A protocol for a systematic review describes the rationale for the review, the objectives, and the methods that will be used to locate, select, and critically appraise studies, and to collect and analyse data from the included studies</td>
</tr>
<tr>
<td>Secondary care</td>
<td>Care provided in hospitals.</td>
</tr>
<tr>
<td>Secondary outcome</td>
<td>An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes</td>
</tr>
<tr>
<td>Selection bias</td>
<td>Selection bias occurs if:</td>
</tr>
<tr>
<td></td>
<td>a) The characteristics of the people selected for a study differ from the wider population from which they have been drawn, or</td>
</tr>
<tr>
<td></td>
<td>b) There are differences between groups of participants in a study in terms of how likely they are to get better</td>
</tr>
<tr>
<td>Selective estrogen-receptor modulator (SERM).</td>
<td>A compound that has a similar chemical structure to oestrogen and has an oestrogen-like effect on some tissues and an anti-oestrogen effect on others</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>How well a test detects the thing it is testing for.</td>
</tr>
<tr>
<td></td>
<td>If a diagnostic test for a disease has high sensitivity, it is likely to pick up all cases of the disease in people who have it (that is, give a 'true positive' result). But if a test is too sensitive it will sometimes also give a positive result in people who don't have the disease (that is, give a 'false positive'). For example, if a test were developed to detect if a woman is 6 months pregnant, a very sensitive test would detect everyone who was 6 months pregnant, but would probably also include those who are 5 and 7 months pregnant. If the same test were more specific (sometimes referred to as having higher specificity), it would detect only those who are 6 months pregnant, and someone who was 5 months pregnant would get a negative result (a 'true negative'). But it would probably also miss some people who were 6 months pregnant (that is, give a 'false negative'). Breast screening is a 'real-life' example. The number of women who are recalled for a second breast screening test is relatively high because the test is very sensitive. If it were made more specific, people who don't have the disease would be less likely to be called back for a second test but more women who have the disease would be missed</td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td>A means of representing uncertainty in the results of an analysis. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. One-way simple sensitivity analysis (univariate analysis): each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. Multi-way simple sensitivity analysis (scenario analysis): 2 or more parameters are varied at the same time and the overall effect on the results is evaluated.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>Term</strong></td>
<td><strong>Definition</strong></td>
</tr>
<tr>
<td>Threshold sensitivity analysis</td>
<td>The critical value of parameters above or below which the conclusions of the study will change are identified. Probabilistic sensitivity analysis: probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (for example, Monte Carlo simulation)</td>
</tr>
<tr>
<td>Sex Steroids</td>
<td>Hormones such as oestrogen, progesterone and testosterone which are produced by the ovaries in women, testes in men, or adrenal gland (in both women and men) that affect the function of the reproductive organs or development of sexual characteristics. Can also be used as medications either in naturally-occurring or synthesised form</td>
</tr>
<tr>
<td>Significance (statistical)</td>
<td>A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 (p&lt;0.05)</td>
</tr>
<tr>
<td>Specificity</td>
<td>The proportion of true negatives that are correctly identified as such. For example in diagnostic testing the specificity is the proportion of non-cases correctly diagnosed as non-cases. See related term ‘Sensitivity’. In terms of literature searching a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers</td>
</tr>
<tr>
<td>Stakeholder</td>
<td>An organisation with an interest in a topic that NICE is developing a clinical guideline or piece of public health guidance on. Organisations that register as stakeholders can comment on the draft scope and the draft guidance. Stakeholders may be: manufacturers of drugs or equipment national patient and carer organisations NHS organisations organisations representing healthcare professionals</td>
</tr>
<tr>
<td>Standard deviation (SD)</td>
<td>A measure of the spread or dispersion of a set of observations, calculated as the average difference from the mean value in the sample</td>
</tr>
<tr>
<td>St. John’s wort</td>
<td>A perennial plant typically used in non-prescription supplement form by some women to ease mild to moderate depression</td>
</tr>
<tr>
<td>Subgroup analysis</td>
<td>An analysis in which the intervention effect is evaluated in a defined subset of the participants in a trial, or in complementary subsets</td>
</tr>
<tr>
<td>Surgical menopause</td>
<td>Induced menopause that results from surgical removal of both of the ovaries in a premenopausal woman. See also Bilateral oophorectomy, Induced menopause</td>
</tr>
<tr>
<td>Systematic review</td>
<td>A review in which evidence from scientific studies has been identified, appraised and synthesised in a methodical way according to predetermined criteria. It may include a meta-analysis</td>
</tr>
<tr>
<td>Tamoxifen</td>
<td>A selective oestrogen-receptor modulator (SERM) that is approved for the prevention and treatment of breast cancer in high-risk women. Although it has an anti-oestrogen effect in the breast, it acts like an oestrogen in the uterus and may cause the lining to thicken. See also Selective oestrogen-receptor modulator</td>
</tr>
<tr>
<td>Testosterone</td>
<td>The male androgen hormone that is essential for sperm production and responsible for inducing and maintaining male secondary sex characteristics. In women, testosterone (partially produced by the ovaries) may regulate sexual desire and may also help maintain bone and muscle health</td>
</tr>
<tr>
<td>Time horizon</td>
<td>The time span over which costs and health outcomes are considered in a decision analysis or economic evaluation</td>
</tr>
<tr>
<td>Treatment allocation</td>
<td>Assigning a participant to a particular arm of a trial</td>
</tr>
<tr>
<td>True negative</td>
<td>A diagnostic test result that correctly indicates that an individual does not</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>True positive</td>
<td>A diagnostic test result that correctly indicates that an individual has the disease of interest, when they do actually have it</td>
</tr>
<tr>
<td>Univariate</td>
<td>Analysis which separately explores each variable in a data set</td>
</tr>
<tr>
<td>Uterus</td>
<td>Womb; small, hollow, pear-shaped organ in a woman’s pelvis where menstrual bleeding originates and in which pregnancy develops. See also Hysterectomy</td>
</tr>
<tr>
<td>Utility</td>
<td>In health economics, a ‘utility’ is the measure of the preference or value that an individual or society places upon a particular health state. It is generally a number between 0 (representing death) and 1 (perfect health). The most widely used measure of benefit in cost-utility analysis is the quality-adjusted life year, but other measures include disability-adjusted life years (DALYs) and healthy year equivalents (HYEs)</td>
</tr>
<tr>
<td>Vagina</td>
<td>The tube that joins the lower part of the uterus to the outside of the body. It is also known as the birth canal</td>
</tr>
<tr>
<td>Vaginal/urogenital atrophy</td>
<td>Thinning and shrinking of the tissues of the vulva, vagina, urethra and bladder caused by lack of oestrogen</td>
</tr>
<tr>
<td>Vaginal dryness</td>
<td>Inadequate lubrication of the vagina, often caused by low oestrogen levels</td>
</tr>
<tr>
<td>Vasomotor symptoms</td>
<td>Menopausal symptoms such as hot flushes and night sweats caused by constriction and dilation of blood vessels that can lead to a sudden increase in blood flow</td>
</tr>
<tr>
<td>Vitamin D</td>
<td>A nutrient that enables the body to absorb calcium, among other things. It is normally produced within the skin in response to sunlight, and absorbed from dietary sources. Also available in supplement form</td>
</tr>
</tbody>
</table>
## 14 Terms, acronyms and abbreviations

<table>
<thead>
<tr>
<th>Terms</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compounded bioidentical hormones</td>
<td>Unregulated plant-derived hormonal combinations similar or identical to human hormones that are compounded by pharmacies to the specification of the prescriber.</td>
</tr>
<tr>
<td>Fragility fracture</td>
<td>Fractures that result from mechanical forces that would not ordinarily result in fracture (such as a fall from a standing height or less). Reduced bone density is a major risk factor for fragility fractures, which occur most commonly in the spine, hip and wrist.</td>
</tr>
<tr>
<td>Low mood</td>
<td>Mild depression symptoms that impair quality of life but are usually intermittent and often associated with hormonal fluctuations in perimenopause.</td>
</tr>
<tr>
<td>Menopause</td>
<td>A biological stage in a woman’s life that occurs when she stops menstruating and reaches the end of her natural reproductive life. Usually it is defined as having occurred when a woman has not had a period for 12 consecutive months (for women reaching menopause naturally). The changes associated with menopause occur when the ovaries stop functioning. Menopause occurs following the cessation of egg (oocyte) maturation and of oestrogen and progesterone secretion.</td>
</tr>
<tr>
<td>Perimenopause</td>
<td>The time in which a woman has irregular cycles of ovulation and menstruation leading up to menopause and continuing until 12 months after her final period (also known as menopausal transition or climacteric).</td>
</tr>
<tr>
<td>Postmenopause</td>
<td>The time after menopause has occurred, starting when a woman has not had a period for 12 consecutive months.</td>
</tr>
<tr>
<td>Premature ovarian insufficiency</td>
<td>Menopause occurring before the age of 40 years (also known as premature ovarian failure or premature menopause). It can occur naturally or as a result of medical or surgical treatment.</td>
</tr>
<tr>
<td>Urogenital atrophy</td>
<td>Thinning and shrinking of the tissues of the vulva, vagina, urethra and bladder caused by oestrogen deficiency that results in multiple symptoms such as vaginal dryness, vaginal irritation, a frequent need to urinate and urinary tract infections.</td>
</tr>
<tr>
<td>Vasomotor symptoms</td>
<td>Menopausal symptoms such as hot flushes and night sweats caused by constriction and dilation of blood vessels in the skin that can lead to a sudden increase in blood flow to allow heat loss.</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
</tbody>
</table>

---

Menopause
Terms, acronyms and abbreviations

National Collaborating Centre for Women’s and Children’s Health
## Menopause

### Terms, acronyms and abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFC</td>
<td>Antral follicle count</td>
</tr>
<tr>
<td>ALP</td>
<td>Alkaline phosphate</td>
</tr>
<tr>
<td>AMH</td>
<td>Anti-Müllerian</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>BKMI</td>
<td>Blatt-Kupperman Menopausal index</td>
</tr>
<tr>
<td>BMD</td>
<td>Bone mineral density</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>BMS</td>
<td>British Medical Society</td>
</tr>
<tr>
<td>BNF</td>
<td>British National Formulary</td>
</tr>
<tr>
<td>CBT</td>
<td>Cognitive behavioural therapy</td>
</tr>
<tr>
<td>CCE</td>
<td>Conjugated equine estrogens</td>
</tr>
<tr>
<td>CEO</td>
<td>Combined equine oestrogens</td>
</tr>
<tr>
<td>CHD</td>
<td>Coronary heart disease</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>DEXA</td>
<td>Dual energy X-ray absorptiometry</td>
</tr>
<tr>
<td>DH</td>
<td>Department of health</td>
</tr>
<tr>
<td>DIC</td>
<td>Deviance information criteria</td>
</tr>
<tr>
<td>DVT</td>
<td>Deep vein thrombosis</td>
</tr>
<tr>
<td>EPT</td>
<td>Oestrogen and progesterone therapy</td>
</tr>
<tr>
<td>ESCIT</td>
<td>Escitalopram</td>
</tr>
<tr>
<td>FRAX</td>
<td>Fracture risk assessment tool</td>
</tr>
<tr>
<td>FSH</td>
<td>Follicle-stimulating hormone</td>
</tr>
<tr>
<td>GDG</td>
<td>Guideline development group</td>
</tr>
<tr>
<td>GNL</td>
<td>GengNianLe</td>
</tr>
<tr>
<td>GP</td>
<td>General practice</td>
</tr>
<tr>
<td>GRADE</td>
<td>Grading of Recommendations Assessment, Development and Evaluation</td>
</tr>
<tr>
<td>HbA1c</td>
<td>Glycated haemoglobin</td>
</tr>
<tr>
<td>HCP</td>
<td>Health care professional</td>
</tr>
<tr>
<td>HDL</td>
<td>High density lipoprotein</td>
</tr>
<tr>
<td>HRT</td>
<td>Hormone replacement therapy</td>
</tr>
<tr>
<td>HT</td>
<td>Hormone therapy</td>
</tr>
<tr>
<td>HTA</td>
<td>Health technology assessment</td>
</tr>
<tr>
<td>ICER</td>
<td>Incremental cost-effectiveness ratio</td>
</tr>
<tr>
<td>IFG</td>
<td>Impaired fasting glycaemia</td>
</tr>
<tr>
<td>IHD</td>
<td>Ischaemic heart disease</td>
</tr>
<tr>
<td>LDL</td>
<td>Low density lipoprotein</td>
</tr>
<tr>
<td>LETR</td>
<td>Linking evidence to recommendations</td>
</tr>
<tr>
<td>LMP</td>
<td>Last menstrual period</td>
</tr>
<tr>
<td>LNG-IUS</td>
<td>Levonorgestrel-secreting intra-uterine system</td>
</tr>
<tr>
<td>MDD</td>
<td>Major depressive disorder</td>
</tr>
<tr>
<td>MHRA</td>
<td>Medicines and healthcare product regulatory authority</td>
</tr>
<tr>
<td>MHT</td>
<td>Menopausal hormone therapy</td>
</tr>
<tr>
<td>MHT</td>
<td>Menopausal hormone therapy</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>MI</td>
<td>Myocardial infarction</td>
</tr>
<tr>
<td>MID</td>
<td>Minimally important difference</td>
</tr>
<tr>
<td>MPA</td>
<td>Medroxyprogesterone acetate</td>
</tr>
<tr>
<td>NCC</td>
<td>National collaborating centre</td>
</tr>
<tr>
<td>NETA</td>
<td>Norethisterone acetate</td>
</tr>
<tr>
<td>NHANES</td>
<td>National Health and Nutrition Examination Survey</td>
</tr>
<tr>
<td>NHS</td>
<td>National health service</td>
</tr>
<tr>
<td>NICE</td>
<td>National institute for health and care excellence</td>
</tr>
<tr>
<td>NIHR</td>
<td>National institute for health research</td>
</tr>
<tr>
<td>NMA</td>
<td>Network meta-analysis</td>
</tr>
<tr>
<td>NPV</td>
<td>Negative predictive value</td>
</tr>
<tr>
<td>OCP</td>
<td>Oral contraceptive pill</td>
</tr>
<tr>
<td>ONS</td>
<td>Office of National Statistics</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>PICO</td>
<td>Population, intervention, comparison, outcome</td>
</tr>
<tr>
<td>POI</td>
<td>Premature ovarian insufficiency</td>
</tr>
<tr>
<td>POI</td>
<td>Premature ovarian insufficiency</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>QALY</td>
<td>Quality adjusted life year</td>
</tr>
<tr>
<td>RCOC</td>
<td>Royal College of Obstetricians and Gynaecologists</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised control trial</td>
</tr>
<tr>
<td>RR</td>
<td>Risk ratio/relative risk</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SE</td>
<td>Standard error</td>
</tr>
<tr>
<td>SNRI</td>
<td>Norepinephrine reuptake inhibitors</td>
</tr>
<tr>
<td>SSRI</td>
<td>Selective serotonin reuptake inhibitors</td>
</tr>
<tr>
<td>STRAW</td>
<td>The Stages of Reproductive Aging Workshop</td>
</tr>
<tr>
<td>SWAN</td>
<td>Study of Women Across the Nation</td>
</tr>
<tr>
<td>T2DM</td>
<td>T2D mellitus</td>
</tr>
<tr>
<td>TS</td>
<td>Turner syndrome</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>VMS</td>
<td>Vasomotor symptoms</td>
</tr>
<tr>
<td>VTE</td>
<td>Venous thromboembolism</td>
</tr>
<tr>
<td>VVA</td>
<td>Vulvovaginal atrophy</td>
</tr>
<tr>
<td>WCH</td>
<td>Women's and Children's health</td>
</tr>
<tr>
<td>WHI</td>
<td>Women's health initiative</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
Appendices

2 The appendices are contained in a separate document:
3 Appendix A: Scope
4 Appendix B: Stakeholders
5 Appendix C: Declarations of interest
6 Appendix D: Review protocols
7 Appendix E: Search strategies
8 Appendix F: Prisma flow charts
9 Appendix G: Excluded studies
10 Appendix H: Evidence tables
11 Appendix I: GRADE profiles
12 Appendix J: Forest plots
13 Appendix K: Network meta-analysis of interventions in the pharmacological and non-pharmacological treatment of short term symptoms for women in menopause
14 Appendix L: Health Economic Analysis