Arch Gynecol Obstet (2011) 284:343–355
DOI 10.1007/s00404-011-1878-x

GENERAL GYNECOLOGY

Hormone therapy in perimenopause and postmenopause (HT)

Interdisciplinary S3 Guideline, Association of the Scientific Medical Societies in Germany AWMF 015/062-short version

Olaf Ortmann · Martina Dören · Eberhard Windler

Received: 11 January 2011 / Accepted: 27 January 2011 / Published online: 24 March 2011

Abstract This short version of the interdisciplinary S3 guideline on hormone therapy in peri- and postmenopause (HT) is intended as a decision-making instrument for physicians and women considering HT. It is designed to assist daily practice. This short version summarises the long version that contains detailed information about the development of the guideline, particularly about establishing the evidence levels. The statements and recommendations, quoted completely, are marked with the relevant levels of evidence (LoE) and grades of recommendation. The classification system from the Centre for Evidence-based Medicine in Oxford was used in this guideline (see "Attachment").

Keywords Hormone therapy · Perimenopause · Postmenopause · Risk communication

O. Ortmann (✉)
Department of Gynaecology and Obstetrics,
University of Regensburg, Landshuter Straße 65,
93053 Regensburg, Germany
e-mail: olaf.ortmann@klinik.uni-regensburg.de
URL: www.caritasstjosef.de

M. Dören
Chair of Women's Health and Osteology,
Charité-Universitätsmedizin Berlin,
Hindenburgdamm 30, 12200 Berlin, Germany
e-mail: epiconsult@t-online.de

E. Windler
Endocrinology and Metabolism of Ageing,
University Medical Center Hamburg-Eppendorf,
Martinistr. 52, Haus N36, 20246 Hamburg, Germany
e-mail: windler@t-online.de
URL: www.studienzentrum-online.de

Abbreviations

HT	Hormone therapy in the perimenopause and postmenopause
ET	Estrogen therapy
EPT	Estrogen–progestin therapy
LoE	Level of evidence
WHI	Women's Health Initiative
HERS	Heart and Estrogen/Progestin Replacement Study
VTE	Venous thromboembolism
CPA	Cyproterone acetate
CMA	Chlormadinon acetate
DNG	Dienogest
DRSP	Drospirenone
WHIMS	The Women's Health Initiative Memory Study
CEE	Conjugated equine estrogens
MPA	Medroxyprogesterone acetate
SSRI	Selective serotonine reuptake inhibitor

Introduction

Peri- and post-menopausal women often seek medical assistance because of climacteric symptoms (e.g. hot flushes and sweating), considering HT for the treatment of these symptoms. They hope for alleviation of these symptoms and possibly an improvement in their quality of life. While ageing, the symptoms may change, and there may be dysfunctions or diseases that also depend on sex hormones. This situation can influence the assessment of the risk–benefit ratio. The prevention of diseases that frequently develop in postmenopausal women is also often an issue when discussing HT.

Compounds, forms of application, pharmacology

There are clinically relevant differences between the available estrogens, progestins and estrogen-progestin combinations as well as between the various forms of application of HT. These differences apply to the benefit as well as to some risks, and they should be taken into account when managing individual cases.

Climacteric symptoms and their treatment

During the climacteric period, women often suffer from vasomotor complaints (hot flushes) and vaginal dryness. These symptoms were the most consistent findings observed in studies in this period of life. Other problems such as disturbed sleep, various physical symptoms, pain, urinary tract disorders, sexual problems and mood fluctuations have a less consistent association with the menopausal transition period. Estrogens are the most effective treatment for hot flushes and vaginal atrophy. The frequency of these complaints can be markedly reduced by HT, or the symptoms will even disappear completely. There are other complaints associated with the climacteric period which may be relieved with HT. If HT is indicated, an improvement of the woman's general well-being is possible.

Statements

Hot flushes and vaginal dryness are associated with the transition from premenopause to postmenopause; the frequency of reports on these symptoms varies (*LoE 2a*).

Other complaints like disturbed sleep, various physical symptoms, urinary tract disorders, sexual problems, or mood fluctuations are inconsistently reported symptoms (*LoE 2a*).

Degree of consensus: strong consensus

Estrogens are effective for the treatment of hot flushes (*LoE 1a*)

Conjugated equine estrogens, oral 17β-estradiol, and transdermal 17β-estradiol reduce hot flushes to a comparable degree (*LoE 1a*).

An additional treatment with progestins does not interfere with the effect of estrogens on vasomotor complaints (*LoE 1a*).

Tibolone is effective in the treatment of hot flushes (*LoE 1a*).

There is no difference in the effect of estrogen therapy on hot flushes between women with natural menopause and women after bilateral oophorectomy (*LoE 1a*).

(see also the chapter on urogenital symptoms).

Degree of consensus: strong consensus

In women who were treated with various estrogens or estrogen-progestin combinations, positive as well as negative effects on the quality of life were found, but in some instances no effect was noted (*LoE 1a*).

Degree of consensus: consensus

Recommendations

When assessing the risk–benefit ratio, one should bear in mind that the only two complaints most consistently reported by women during the menopausal transition are hot flushes and vaginal dryness (*A*).

Degree of consensus: consensus

Hot flushes can be treated with estrogens or, if applicable, estrogen-progestin combinations, or with tibolone (*A*).

When deciding on the indication, the risks and benefits presented in this guideline should be taken into account (*A*).

Degree of consensus: strong consensus

The sole improvement of a woman's so-called general well-being or health-related quality of life is not an indication for HT (B).

Degree of consensus: strong consensus

Vulvovaginal atrophy

Systemic HT or local estrogen therapy (ET) prevents or reverses vaginal atrophy. Low-dose local ET is as effective as systemic treatment. Local ET is significantly more effective than placebo or local therapy without hormones.

Statement

HT is effective for preventing and/or treating vaginal atrophy (*LoE 1a*).

Degree of consensus: strong consensus

Recommendation

If symptomatic vaginal atrophy is the only indication for therapy, local vaginal ET should be recommended (*A*).

Degree of consensus: consensus

Urinary incontinence

Earlier studies reported that ET can improve or even cure urinary incontinence, especially urge incontinence. Including data from the Women's Health Initiative (WHI) study and the Heart and Estrogen/progestin Replacement Study (HERS), more recent systematic reviews concluded that oral HT increases the risk for urinary incontinence or may lead to a deterioration of existing incontinence. Transdermal or vaginal estrogen application resulted in inconclusive improvement of incontinence. The efficacy of

various physical, surgical and non-hormonal treatments has been proven. Local ET is often used as an adjunct combined with i.e. surgical therapies.

Statement

Oral HT has a negative influence on urinary incontinence (*LoE 1a*).

A clear positive effect of local and transdermal therapies could not be demonstrated (*LoE 1a*).

Degree of consensus: consensus

Recommendation

Oral HT should not be recommended for the treatment of urinary incontinence (*B*).

Degree of consensus: consensus

Statement

There are other medications and therapeutic measures with proven efficacy for the treatment of urinary incontinence that should be recommended (*LoE 1a*).

Degree of consensus: consensus

Recurrent urinary tract infection

Studies with oral HT did not show any effect on the incidence of urinary tract infections. In some smaller studies and studies with heterogeneous methodologies, vaginal estrogen treatment led to a significant reduction of urinary tract infections.

Statement

Oral HT is not appropriate for the prevention of recurrent urinary tract infections (*LoE 1a*).

Vaginal estrogen therapy is effective (*LoE 2a*).

Degree of consensus: strong consensus

Recommendation

Vaginal estrogen treatment can be recommended for recurrent urinary tract infections (*B*).

Degree of consensus: strong consensus

Locomotor system and bone metabolism

A reduction of the incidence of bone fractures by HT was shown in a large number of studies. HT reduced the rate of clinical fractures as well as the rate of osteoporosis-associated fractures. Even low doses (0.3 mg of conjugated estrogens, 0.5 mg of oral 17-β-estradiol or 14 μg of transdermal 17-β-estradiol) have been shown to reduce the loss of bone mass; a reduction in fracture rates has not been proven conclusively with these doses. HT is effective in the primary prevention of osteoporosis and osteoporosis-associated fractures.

Statement

HT significantly reduces the incidence of fractures (*LoE 1a*).

Degree of consensus: strong consensus

Recommendation

After consideration of risks and benefits, HT can be used for the prevention of fractures, for those women at high risk, if the woman cannot tolerate other first-line drugs recommended for the treatment of osteoporosis or if these drugs are contraindicated (*A*).

Degree of consensus: consensus

Coronary heart disease

HT is not indicated for the primary or secondary prevention of coronary heart disease in women at any age because there are other strategies with proven efficacy. However, numerous findings from observational studies suggest that HT may reduce the risk of myocardial infarction, when started early. In contrast to these findings, the WHI study, a randomized controlled trial, found a non-significant trend of risk reduction by ET in women aged 50–59 years, but not in older women. With combined estrogen–progestin therapy (EPT), the risk was increased at the beginning of treatment, but not after treatment duration of 5.6 years. Particularly, older women and women with a predisposition for cardiovascular disease had an increased risk at the beginning of treatment. Especially in older women (>60 years), systemic HT should only be initiated after careful assessment of risks and benefits, taking any pre-existing risk factors into consideration.

Recommendation

HT is not indicated for the primary or secondary prevention of coronary heart disease (*B*).

There are other strategies with proven efficacy for primary or secondary prevention (*A*).

Degree of consensus: strong consensus

Cerebrovascular disease

HT increases the risk for an ischaemic cerebrovascular event. This risk should always be considered in the assessment of risks and benefits, particularly in older women.

Statement

HT increases the risk for an ischaemic cerebrovascular event (*LoE 1a*).

Degree of consensus: strong consensus

Recommendation

The increased risk of stroke must always be considered in the assessment of risks and benefits of HT (*A*).

Degree of consensus: strong consensus

Thromboembolic disease

HT increases the risk of venous thrombosis and pulmonary embolism, particularly in the first year of use and in the presence of risk factors such as congenital coagulation disorders. A meta-analysis of observational studies shows that the risk is lower with transdermal HT.

Statement

Oral HT increases the risk of venous thrombosis and pulmonary embolism (VTE) (*LoE 1a*).

Degree of consensus: strong consensus

Recommendation

The increased risk of VTE must be considered while assessing the risks and benefits of HT. The risk is particularly high in the first year of HT use and further increases if there are additional predisposing risk factors for venous thrombosis (*A*).

Degree of consensus: strong consensus

Skin ageing

Available data are insufficient to make reliable statements on the effects of HT on the processes of skin ageing. Smaller comparative studies have, in part, shown a positive effect of estrogens on parameters of skin ageing. The limited number of randomized studies with small numbers of cases and significant methodological flaws, respectively have not yielded any reliable results.

Statement

An alleviation of skin ageing processes by HT has not been proven (*LoE 2b*).

Degree of consensus: strong consensus

Recommendation

HT is not indicated for alleviating the processes of skin ageing (*A*).

Degree of consensus: strong consensus

Signs of skin androgenisation

There is only a small number of evaluable trials using HT with anti-androgenic progestins (cyproterone acetate [CPA], chlormadinone acetate [CMA], dienogest [DNG], and drospirenone [DRSP]) in the climacteric period. In particular, it is impossible to state if HT with anti-androgenic progestins can result in a significant improvement of androgenic skin changes, because there have been no specific trials addressing this question. However, if a combined estrogen–progestin therapy is indicated, women with cutaneous signs of androgenisation should primarily receive preparations with an anti-androgenic progestin compound instead of a preparation with a progestin deriving from the 19-nortestosterone group.

Statement

Alleviation of the signs of skin androgenisation with HT has not been proven (*LoE 5*).

Degree of consensus: strong consensus

Recommendation

HT is not indicated to alleviate signs of skin androgenisation (*A*).

Degree of consensus: strong consensus

Diseases of the gall bladder and gall ducts

HT increases the risk of gall duct disease. This is mainly the effect of the estrogen compound. The risk is probably less profound with transdermal estrogen application.

Statement

There is evidence that diseases of the gall bladder and gall ducts, particularly cholecystolithiasis, cholecystitis/

cholangitis as well as cholecystectomies, occur more frequently with HT (*LoE Ib*).

Degree of consensus: strong consensus

Recommendation

When assessing the risks and benefits of HT, the increased risk of cholecystitis/cholangitis, cholecystolithiasis and cholecystectomies must be taken into account (*A*).

Degree of consensus: consensus

Cognition

There is limited evidence from older clinical trials that ET has a short-term positive effect on cognition when used in premenopausal women after bilateral oophorectomy. The long-term effects of HT started during the menopausal transition or during the early postmenopausal period are unknown. Neither ET nor EPT were able to prevent the decline of cognitive functions in older postmenopausal women, either as short- or long-term treatment. The evidence is insufficient to assess if special forms of HT may confer any benefit.

Statement

HT does not have a positive effect on cognition in older postmenopausal women (*LoE 2a*).

Degree of consensus: strong consensus

Recommendation

HT should not be recommended to alleviate impairments of cognition in postmenopausal women (*B*).

Degree of consensus: strong consensus

Dementia

Observational studies have demonstrated a reduction of the risk for dementia, e.g. Alzheimer's disease, with the use of HT. However, these studies are heterogeneous and show a substantial bias. Therefore, based on the insufficient quality of the data, it is not possible to make recommendations on the basis of available evidence.

In the Women's Health Initiative Memory Study (WHIMS), the relationship between HT and dementia in women over 65 years was investigated as part of the WHI. The endpoint "mild cognitive impairment" did not show any difference between HT and placebo, either with conjugated equine estrogens (CEE) plus medroxyprogesterone acetate (MPA), or with CEE alone. With the endpoint "possible dementia", there was a significantly increased relative risk for the combination of CEE and MPA, but not for CEE alone, versus placebo.

In women with a diagnosis from mild to moderate Alzheimer's dementia, there was no significant difference between 1-year ET and placebo as to the overall appearance of the Alzheimer's dementia.

Statements

HT does not show any benefit on the signs of dementia in women with Alzheimer's disease (*LoE 1a*).

Degree of consensus: strong consensus

Continuous combined HT increases the risk of dementia in women aged over 65 years (*LoE 2a*).

Degree of consensus: strong consensus

Recommendation

HT should not be recommended to decrease the risk of dementia (*A*).

Degree of consensus: strong consensus

Breast cancer

The use of HT increases the risk of breast cancer. The increased risk was seen after duration of use of 5 or more years. Meta-analyses incorporating both observational studies and randomized controlled trials have also shown an increased risk of breast cancer with ET alone. The effect was weaker than with EPT. In addition, the increase of risk by ET, compared with EPT, was only seen after a longer duration of use. The WHI did not show any increased risk after a mean duration of ET use of 7.1 years. After HT is discontinued, the risk decreases. After a few years, the risk does not differ from the risk of women who never used HT.

Statements

EPT increases the risk of breast cancer (*LoE 1b*).

ET increases the risk of breast cancer to a lesser degree than EPT (*LoE 2a*).

Degree of consensus: strong consensus/consensus

Recommendation

The increased risk of breast cancer must be considered while assessing the benefits and risks of HT (*A*).

Degree of consensus: strong consensus

Endometrial cancer

ET leads to an increased risk of endometrial cancer. The effect is dependent on the duration of use and on the estrogen dose. EPT combined with a progestin, applied at an adequate dosage for at least 10 days per treatment month, does not increase the risk of endometrial cancer.

Low-dose vaginal ET, as used to prevent vaginal atrophy, probably does not increase the risk of endometrial cancer. However, data on this issue are very limited.

Statement

ET increases the risk of endometrial cancer, whereas combined EPT with at least ten; better twelve days of progestin application per treatment month does not. (*LoE 1a*).

Degree of consensus: strong consensus

Recommendation

ET should only be used in hysterectomized women. Combined EPT in non-hysterectomized women should include a progestin application for at least 10 days, better 12 days per treatment month (*A*).

Degree of consensus: strong consensus

Ovarian cancer

In the past, the relation between HT use and the risk of ovarian cancer was discussed controversially. Recent meta-analyses have shown an increased risk for ovarian cancer with the use of ET or EPT.

Statement

HT increases the risk of ovarian cancer. It is unclear if there are any differences between ET and EPT (*LoE 2a*).

Degree of consensus: strong consensus

Recommendation

The increased risk of ovarian cancer must be considered while assessing the risks and benefits of HT (*A*).

Degree of consensus: strong consensus

Colorectal cancer

According to observational studies, the risk of colorectal cancer is reduced in women who have used ET or EPT. The risk reduction was more pronounced in current users of HT. In the WHI, a randomized, placebo-controlled trial, only EPT led to a significant risk reduction.

Statement

EPT decreases the risk of colorectal cancer; ET does not (*LoE 2a*).

Degree of consensus: consensus

Recommendation

This does not result in an indication for HT (*A*).

Degree of consensus: strong consensus

HT in cancer patients

According to a recent randomized trial, HT use after breast cancer leads to a markedly increased risk of recurrence.

Assessing the risk of HT after endometrial, ovarian or colorectal cancer is difficult because only few observational studies are available. These did not demonstrate an increased risk of recurrence when HT was used. Yet, the number of cases is too low to draw reliable conclusions about the safety of HT after treatment of the cancer entities mentioned earlier.

Statements

HT increases the risk of recurrence when used after breast cancer (*LoE 2b*).

The risk of HT use after treated endometrial, ovarian or colorectal cancer has not been studied sufficiently (*LoE 2b*).

Since lack of data, no statement is possible regarding other types of tumours (*LoE 5*).

Degree of consensus: strong consensus/consensus

Recommendation

HT is contraindicated after breast cancer treatment (*A*).

Degree of consensus: strong consensus

Premature menopause

Women with premature menopause (<40 years of age) are a heterogeneous group. The available studies have mostly investigated women with surgical oophorectomy. From a clinical point of view, it seems to make sense to use HT in women with premature menopause, at least until the average age of natural menopause (about 50 years).

Statements

It is unclear whether the benefits and risks of HT in women with premature menopause differ from those in women with natural menopause aged around 50 years (*LoE 2a*).

HT is appropriate for the treatment of hot flushes and vaginal atrophy in symptomatic women with premature menopause (*LoE 2a*).

Degree of consensus: strong consensus/consensus

Recommendation

In women with premature menopause, HT can be used until the average age of natural menopause (0).

Degree of consensus: consensus

Alternative therapies

Currently, there is no evidence that herbal remedies have a reliable effect on vasomotor complaints. Isoflavones or *Cimicifuga racemosa* can be considered in cases of mild hot flushes or sweating, since a decrease of climacteric complaints is possible in few cases. It is not possible to predict if this treatment will be effective in the individual case. In cases of severe vasomotor complaints, a sufficient therapeutic effect cannot be expected. If there are contraindications against hormonal therapies, and the woman expresses an urgent desire for treatment, selective serotonin re-uptake inhibitors (SSRI) and gabapentin may be considered as an individual experimental treatment. It should be noted that neither medication is currently approved for this indication. Therefore, a medical justification based on the risk–benefit assessment is needed, and the patient must be thoroughly informed about the situation ("off label use"). For all alternative therapies, there is a lack of data on long-term safety.

Statements

Isoflavone containing supplements made from soybeans or red clover, or a nutrition rich in phytoestrogens, do not reduce hot flushes, or only marginally, if at all (*LoE 1a*).

Currently, the possible risks of alternative therapies cannot be assessed with sufficient reliability (*LoE 1a*).

Degree of consensus: strong consensus

Recommendation

Phytoestrogens and other herbal or non-hormonal therapies cannot be recommended as an alternative to HT (0).

Degree of consensus: strong consensus

Risk communication

Risk communication is defined as communication about the probabilities of expected benefits and the possible risks of harm by HT, with the patient and possibly with an accompanying person.

For an individual assessment and evaluation of the probability of benefit and the risk of harm, individual factors such as the woman's general state of health, age, age at menopause, previous HT, duration of use, dosage and type of HT, and diseases while using HT should be taken into consideration. In order to give adequate information about the risks to the woman seeking advice, the doctor must be familiar with the principles of risk calculation. He or she should also be able to communicate the probabilities in such a way that the patient can make her own individual decision for or against the initiation of HT. The figures necessary for this communication can be found in the long version and in the balance sheet (see "attachments").

Attachment

Levels of evidence

Oxford Centre for Evidence-based Medicine Levels of Evidence. (March 2009) (for definitions of terms used see glossary at http://www.cebm.net/?o=1116).

Level	Therapy/ prevention, aetiology/harm	Prognosis	Diagnosis	Differential diagnosis/symptom prevalence study	Economic and decision analyses
1a	SR (with homogeneity*) of RCTs	SR (with homogeneity*) of inception cohort studies; CDR† validated in different populations	SR (with homogeneity*) of Level 1 diagnostic studies; CDR† with 1b studies from different clinical centres	SR (with homogeneity*) of prospective cohort studies	SR (with homogeneity*) of Level 1 economic studies

Table continued

Level	Therapy/ prevention, aetiology/harm	Prognosis	Diagnosis	Differential diagnosis/ symptom prevalence study	Economic and decision analyses
1b	Individual RCT (with narrow confidence interval[‡])	Individual inception cohort study with >80% follow-up; CDR[†] validated in a single population	Validating** cohort study with good[†††] reference standards; or CDR[†] tested within one clinical centre	Prospective cohort study with good follow-up****	Analysis based on clinically sensible costs or alternatives; systematic review(s) of the evidence; and including multi-way sensitivity analyses
1c	All or none[§]	All or none case-series	Absolute SpPins and SnNouts[††]	All or none case-series	Absolute better-value or worse-value analyses[††††]
2a	SR (with homogeneity*) of cohort studies	SR (with homogeneity*) of either retrospective cohort studies or untreated control groups in RCTs	SR (with homogeneity*) of Level >2 diagnostic studies	SR (with homogeneity*) of 2b and better studies	SR (with homogeneity*) of Level >2 economic studies
2b	Individual cohort study (including low quality RCT; e.g. <80% follow-up)	Retrospective cohort study or follow-up of untreated control patients in an RCT; Derivation of CDR[†] or validated on split-sample[§§§] only	Exploratory** cohort study with good[†††] reference standards; CDR[†] after derivation, or validated only on split-sample[§§§] or databases	Retrospective cohort study, or poor follow-up	Analysis based on clinically sensible costs or alternatives; limited review(s) of the evidence, or single studies; and including multi-way sensitivity analyses
2c	"Outcomes" research; ecological studies	"Outcomes" research		Ecological studies	Audit or outcomes research
3a	SR (with homogeneity*) of case–control studies		SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies
3b	Individual case–control study		Non-consecutive study; or without consistently applied reference standards	Non-consecutive cohort study, or very limited population	Analysis based on limited alternatives or costs, poor quality estimates of data, but including sensitivity analyses incorporating clinically sensible variations
4	Case-series (and poor quality cohort and case–control studies[§§])	Case-series (and poor quality prognostic cohort studies***)	Case–control study, poor or non-independent reference standard	Case-series or superseded reference standards	Analysis with no sensitivity analysis

Table continued

Level	Therapy/prevention, aetiology/harm	Prognosis	Diagnosis	Differential diagnosis/ symptom prevalence study	Economic and decision analyses
5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on economic theory or "first principles"

Produced by Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Dawes since November 1998. Updated by Jeremy Howick March 2009

Notes: Users can add a minus-sign "−" to denote the level of that fails to provide a conclusive answer because: either a single result with a wide confidence interval or a systematic review with troublesome heterogeneity. Such evidence is inconclusive, and therefore can only generate Grade D recommendations

* By homogeneity we mean a systematic review that is free of worrisome variations (heterogeneity) in the directions and degrees of results between individual studies. Not all systematic reviews with statistically significant heterogeneity need be worrisome, and not all worrisome heterogeneity need be statistically significant. As noted above, studies displaying worrisome heterogeneity should be tagged with a "−" at the end of their designated level

† Clinical Decision Rule. (These are algorithms or scoring systems that lead to a prognostic estimation or a diagnostic category.)

‡ See note above for advice on how to understand, rate and use trials or other studies with wide confidence intervals

§ Met when all patients died before the Rx became available, but some now survive on it; or when some patients died before the Rx became available, but none now die on it

§§ By poor quality cohort study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably blinded), objective way in both exposed and non-exposed individuals and/or failed to identify or appropriately control known confounders and/or failed to carry out a sufficiently long and complete follow-up of patients. By poor quality case–control study we mean one that failed to clearly define comparison groups and/or failed to measure exposures and outcomes in the same (preferably blinded), objective way in both cases and controls and/or failed to identify or appropriately control known confounders

§§§ Split-sample validation is achieved by collecting all the information in a single tranche, then artificially dividing this into "derivation" and "validation" samples

†† An "Absolute SpPin" is a diagnostic finding whose specificity is so high that a positive result rules-in the diagnosis. An "Absolute SnNout" is a diagnostic finding whose sensitivity is so high that a negative result rules-out the diagnosis

‡‡ Good, better, bad and worse refer to the comparisons between treatments in terms of their clinical risks and benefits

††† Good reference standards are independent of the test, and applied blindly or objectively to applied to all patients. Poor reference standards are haphazardly applied, but still independent of the test. Use of a non-independent reference standard (where the 'test' is included in the 'reference', or where the 'testing' affects the 'reference') implies a level 4 study

†††† Better-value treatments are clearly as good but cheaper, or better at the same or reduced cost. Worse-value treatments are as good and more expensive, or worse and the equally or more expensive

** Validating studies test the quality of a specific diagnostic test, based on prior evidence. An exploratory study collects information and trawls the data (e.g. using a regression analysis) to find which factors are 'significant'

*** By poor quality prognostic cohort study we mean one in which sampling was biased in favour of patients who already had the target outcome, or the measurement of outcomes was accomplished in <80% of study patients, or outcomes were determined in an unblinded, non-objective way, or there was no correction for confounding factors

**** Good follow-up in a differential diagnosis study is >80%, with adequate time for alternative diagnoses to emerge (for example 1–6 months acute, 1–5 years chronic)

Grades of recommendation

A	Consistent level 1 studies
B	Consistent level 2 or 3 studies or extrapolations from level 1 studies
C	Level 4 studies or extrapolations from level 2 or 3 studies
D	Level 5 evidence or troublingly inconsistent or inconclusive studies of any level

"Extrapolations" are where data is used in a situation that has potentially clinically important differences than the original study situation

Consensus criteria for the degree of recommendation:	
Consistency of trial results	
Clinical relevance of endpoints and potency of effect	
Risk–benefit ratio	
Ethical commitment	
Patient preference	
Feasibility, practicability	
Degree	Degree of recommendation for the option for action according to (2)
A	Strong recommendation 'shall'
B	Recommendation 'should'
0	Recommendation open 'can' (option for action)
Negative recommendations are expressed by using the word 'not' with the same symbols	
GA	Guideline adaptation
GCP	Good clinical practice
GEP	Good epidemiology practice

Balance sheet

Endpoint (EPT)	Relative risks (RR) ET: estrogen therapy, EPT: estrogen-progestin therapy	Absolute risks (AR)	Number needed to harm (+NNH)/number needed to treat (−NNT)
Hot flushes	OR 0.13 (95% CI 0.07–0.23)	n/a	n/a
Recurrent urinary tract infections	Vaginal ET (2 studies):	n/a	n/a
	RR 0.25 (95% CI 0.13–0.30)		
	RR 0.64 (95% CI 0.47–0.86)		
Coronary heart disease	ET: myocardial infarction and coronary death: HR 0.91 (95% CI 0.75–1.12)	−5 events/10,000 women/year of use (corresponding to 49 events [hormone group] versus 54 events [placebo group]; statistically not significant)	
	After myocardial infarction: HR 0.99 (95% CI 0.70–1.41)		
	EPT: HR 1.24 (95% CI 1.00–1.54)	+6 events/10,000 women/year of use (39 events [hormone group] versus 33 events [placebo group])	+1,667

Table continued

Endpoint (EPT)	Relative risks (RR) ET: estrogen therapy, EPT: estrogen-progestin therapy	Absolute risks (AR)	Number needed to harm (+NNH)/number needed to treat (−NNT)
Stroke	ET: cerebrovascular accident: HR 1.39 (95% CI 1.10–1.77)	+12 events/10,000 women/year of use (44 events [hormone group] versus 32 events [placebo group])	+833
	EPT: ischaemic stroke: HR: 1.44 (95% CI 1.09–1.90)	+8 events/10,000 women/year of use (26 events [hormone group] versus 18 events [placebo group])	+1,250
	Haemorrhagic stroke: HR 0.82 (95% CI 0.43–1.56)	+0 events/10,000 women/year of use (4 events [hormone group] versus 4 events [placebo group])	
Thromboembolic events	ET: HR 1.47 (95% CI adjusted 0.87–2.47)	+6 events/10,000 women/year of use (21 events [hormone group] versus 15 events [placebo group])	+1,667
	EPT: HR 2.06 (95% CI adjusted 1.57–2.70)	+17 events/10,000 women/year of use (35 events [hormone group] versus 17 events [placebo group])	+588
Dementia	EPT: RR 1.97 (95% CI 1.16–3.33)	+23 events/10,000 women/year of use (45 events [hormone group] versus 22 events [placebo group])	+435
Fractures	EPT: femoral neck fractures: HR 0.66 (95% CI 0.45–0.98)	−5 events/10,000 women/year of use (10 fractures [hormone group] versus 15 fractures [placebo group])	–2,000
	Vertebral body fractures: HR 0.66 (95% CI 0.44–0.98)	−6 events/10,000 women/year of use (9 fractures [hormone group] versus 15 fractures [placebo group])	–1,667
	Total rate of fractures: HR 0.76 (95% CI 0.69–0.85)	−44 events/10,000 women/year of use (147 fractures [hormone group] versus 191 fractures [placebo group])	–227
	ET: fractures of the proximal femur: HR 0.61 (95% CI 0.41–0.91)	−6 events/10,000 women/year of use (11 fractures [hormone group] versus 17 fractures [placebo group])	–1,667
	Vertebral body fractures: HR 0.62 (95% CI 0.42–0.93)	−6 events/10,000 women/year of use (11 [hormone group] versus 17 fractures [placebo group])	−1,667
	Total rate of fractures: HR 0.70 (95% CI 0.63–0.79)	−56 events/10,000 women/year of use (139 fractures [hormone group] versus 195 fractures [placebo group])	−179
Gall duct disease (any)	ET: HR 1.67 (95% CI 1.35–2.06)	+31 events/10,000 women/year of use (78 events [hormone group] versus 47 events [placebo group])	+323
	EPT: HR 1.59 (95% CI 1.28–1.97)	+20 events/10,000 women/year of use (55 [hormone group] versus 35 events [placebo group])	+500
Breast cancer	EPT: RR 1.26 (95% CI 1.00–1.59)	+8 breast cancers/10,000 women/year of use (38 events [hormone group] versus 30 events [placebo group])	+1,250
	ET: RR 0.77 (95% CI 0.59–1.01)	−7 breast cancers/10,000 women/year of use (statistically non-significant)	
Ovarian cancer	EPT: RR 1.11 (95% CI 1.020–1.207)		
	ET: RR 1.284 (95% CI 1.178–1.399)		
Colorectal cancer	EPT: HR 0.63 (95% CI 0.43–0.92)	−6 colorectal carcinomas/10,000 women/year of use (10 events [hormone group] versus 16 events [placebo group])	−1,667
	ET: HR 1.08 (95% CI 0.75–1.55)	+1 colorectal carcinoma/ 10,000 women/year of use (statistically non-significant)	
After breast cancer	EPT: HR 2.4 (95% CI 1.3–4.2)		

For references, see the chapter "Risk Communication" in the long version

The balance sheet listed above is intended for demonstration of the risk of HT with respect to different endpoints.

There are different indices available to quantify the effect of interventions:

Absolute risk reduction (ARR) describes the absolute difference in the rate of undesirable events in the experimental group (E) compared with the control group (C) if the experimental treatment is effective (ARR = C − E).

The reciprocal value of the ARR is the number needed to treat (1/ARR = NNT). NNT is a clinically intuitive measure for endpoints to describe the effects of a certain treatment. It represents the number of patients who must be treated to prevent one additional undesirable event.

The absolute risk increase (ARI) describes the absolute difference in the rate of undesirable events in the experimental group in comparison with the control group if the experimental treatment is worse (ARI = |C − E|).

The reciprocal value of the ARI is the number needed to harm (NNH). NNH is a clinically intuitive measure for endpoints to describe the unwanted effects of a certain treatment. It represents the number of patients who must be treated to cause one additional undesirable event.

The relative risk reduction (RRR) describes the relative decrease in the rate of undesirable events in the experimental study group as compared with the control group (RRR = |C − E|/C).

Example: Phlebothrombosis

If the yearly rate of phlebothrombosis in postmenopausal users of oral ET is 22 in 10,000 women, and the rate in non-users is 11 in 10,000 women, the RR is

$$\mathrm{RR} = \frac{22}{10,000/\mathrm{year}} \div \frac{11}{10,000/\mathrm{year}} = 2$$

This means a doubled risk of phlebothrombosis when using ET for 1 year. An RR of more than 1.0 indicates an increased risk. An RR of 1.2 signifies a risk increase by 20%. An RR of less than 1.0 indicates a decrease of risk. An RR of 0.5, for example, would mean a risk decrease by 50%—the probability of an event when using ET would then be only half as high as with non-use.

For the evaluation of risks, it is often more useful to state the absolute risk (AR). The AR describes the risk difference by calculating the difference in the incidence between exposed and non-exposed populations. In the example used above (phlebothrombosis in ET users) the AR is

$$\mathrm{AR} = \frac{22}{10,000/\mathrm{year}} - \frac{11}{10,000/\mathrm{year}} = \frac{11}{10,000/\mathrm{year}}$$

This means that there will be 11 additional phlebothromboses per year for every 10,000 women using oral ET. Changes in the AR are, however, significantly influenced by the pre-existing risk found in the exposed persons.

Report on guidelines and methods

The long version, the list of references and the detailed report on guidelines and methods are published in German on the DGGG homepage (http://www.dggg.de, area "Leitlinien").

This guideline was developed under the aegis of the DGGG and approved by the following institutions:

Medical associations

Deutsche Gesellschaft für Gynäkologie und Geburtshilfe (DGGG).

Deutsche Krebsgesellschaft (DKG).

Deutsche Gesellschaft für Innere Medizin (DGIM).

Deutsche Gesellschaft für Endokrinologie (DGE).

Deutsche Gesellschaft für Kardiologie (DGK).

Deutsche Gesellschaft für Neurologie (DGN).

Deutsche Gesellschaft für Senologie (DGS).

Dachverband Osteologie (DVO).

Deutsche Gesellschaft für Angiologie (DGA).

Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS).

Deutsche Menopause Gesellschaft (DMG).

Deutsche Gesellschaft für Psychosomatische Frauenheilkunde und Geburtshilfe (DGPFG).

Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM).

Deutsche Gesellschaft für Gynäkologische Endokrinologie und Fortpflanzungsmedizin (DGGEF).

Working groups

Arbeitsgemeinschaft Gynäkologische Onkologie (AGO).

Arbeitsgemeinschaft für Urogynäkologie und plastische Beckenbodenrekonstruktion (AGUB).

Professional organisations

Berufsverband für Frauenärzte (BVF).

Other institutions

Women's Health Coalition (WHC).

Frauenselbsthilfe nach Krebs e.V.

Contributors

C. Albring, E. Baum, M.J. Beckermann, M.W. Beckmann, M. Blettner, B. Böhm, C. Brucker, M. Dören, G. Emons, D. Foth, F. Geisthövel, Th. Gudermann, P. Hadji, L. Kiesel, D. Klemperer, K. König, E. Lindhoff-Last, A. Ludolph, A.O. Mueck, I. Naß-Griegoleit, D. Noss, O. Ortmann, E. Petri, Th. Rabe, V. Regitz-Zagrosek, H. Schulte, F. Siedentopf, Th. Strowitzki, E. Windler.

AWMF 015/062

Valid until December 2011.

Acknowledgments The authors are grateful to F. Günther and S. Block for the preparation of the English version of the guideline.

Conflict of interest We declare that we have no conflict of interest.