

# Selection of Key Recommendations for the Management of Women with Endometriosis by an International Panel of Patients and Professionals

Schleedoorn MJ<sup>1,\*</sup>, Nelen WLDM<sup>1</sup>, Dunselman GAJ<sup>2</sup>, Vermeulen N<sup>3</sup>, on behalf of the EndoKey Group<sup>\*\*</sup>.

## Author Affiliations

<sup>1</sup>Department of Obstetrics and Gynaecology, Radboud University Medical Center, PO Box 9101, 6500HB Nijmegen, The Netherlands

<sup>2</sup> Department of Obstetrics and Gynaecology, Research Institute GROW, Maastricht University Medical Center, PO Box 5800, 6202AZ Maastricht, The Netherlands

<sup>3</sup> European Society of Human Reproduction and Embryology, Central Office, Meerstraat 60, 1852 Grimbergen, Belgium

\*Correspondence address. E-mail: [myra.schleedoorn@radboudumc.nl](mailto:myra.schleedoorn@radboudumc.nl)

\*\* Members of the EndoKey Group are listed in the Acknowledgements

## Abstract

**Study question:** Is it possible to achieve consensus between patients and professionals on one set of key recommendations and are there differences in patients' perspective and professionals' perspective regarding essential endometriosis care?

**Summary answer:** Patients and professionals reached consensus on one set of key recommendations. This set covers all fields of endometriosis care. Significant differences between patient's perspective and professional's perspective were seen in the first round but disappeared in the following 2 rounds when opinions were swayed.

**What is known already:** Guideline development alone will not lead to health care improvement. Quality indicators are needed to monitor actual care and guideline adherence. These can help to better implement the ESHRE guideline in European hospitals and thereby improve the quality of

endometriosis care. The first step in the development of quality indicators is to select a compact set of key recommendations.

**Study design, size and duration:** Using a RAND modified Delphi method, this study reports the systematic selection of key recommendations based on the ESHRE guideline Management of Women with Endometriosis by an international expert panel of both patients and professionals during the study period of September 2015 and December 2015.

**Participants, setting, methods:** An international panel of patients (n=10) and medical professionals (n=11) rated and prioritized the 83 recommendations extracted from the ESHRE Guideline for relevance in 3 rounds. A strict consensus methodology was used to select key recommendations. The main outcome measure was one set of key recommendations for endometriosis care.

**Main results and the role of chance:** A representative set of 17 key recommendations was selected from the preliminary set of 83 recommendations. This selection covers all dimensions of endometriosis care, including diagnosis, treatment of endometriosis-associated pain, treatment of endometriosis-associated infertility and miscellaneous topics as prevention, menopause and relationship with cancer. 16 (76,2%) of the 21 experts participated in all three rounds.

**Limitations, reasons for caution:** The feasibility of the selected key recommendations was not assessed in this study. As not all panel members took part in all three rounds, some response bias may have occurred.

**Wider implications of the findings:** This set of key recommendations is the first step in the development of quality indicators for monitoring and improving endometriosis care. The set is

generic and can be used in hospitals internationally. A practice test should be conducted to assess the feasibility of our key recommendations in clinical practice.

**Study funding / competing interest(s):** Members of the EndoKey study group did not receive payment. Competing interests: none.

**Key words:** *endometriosis / key recommendations / quality indicators / quality of care / European Society of Human Reproduction and Embryology*

## Introduction

Endometriosis is defined as the presence of endometrial glands and stroma outside the uterus where it causes a chronic inflammatory reaction (Kennedy et al. 2005). It is one of the most common gynaecological disorders with an estimated prevalence of 2-10% within the worldwide female population (Eskenazi and Warner 1997). However, the prevalence may go up to 50% in infertile women (Meuleman et al. 2009). Endometriosis can affect women of reproductive age, causing infertility, pain and can have a significant negative impact on different psychosocial aspects of a woman's life (Moradi et al. 2014; de Graaff et al. 2013; Culley et al. 2013; Nnoaham et al. 2011). Since there is currently no cure, treatment focuses on reducing endometriosis-associated pain and improving fertility.

Many medical professionals experience difficulties in the management of women with endometriosis, which is reflected in the wide variety of clinical practice among European countries (Johnson and Hummelshoj 2013). As a result, many patients receive either delayed or suboptimal care (Kennedy et al. 2005; Ballard, Lowton, and Wright 2006). The World Endometriosis Research Foundation (WERF)'s EndoCost study estimates the total costs arising from women with endometriosis between 0.8 million–12.5 billion euro per European country per year (Simoens et al. 2012). The total annual costs, including indirect costs of productivity loss related to endometriosis are estimated at €9,872 (95% CI €7,930 – 11,870) per patient, with costs of productivity loss representing 75% of total costs (Klein et al. 2014). The direct annual healthcare costs of €2,238 (95% CI €1,567 – 3,240) per patient suffering from endometriosis are similar to those of diabetes mellitus (€2,858) (von Ferber, Koster, and Hauner 2006).

The ESHRE Guideline *Management of Women with Endometriosis* (Dunselman et al. 2014) aimed to improve European endometriosis care by providing 83 recommendations based on literature evidence and good clinical practice. Unfortunately, guideline development is not automatically followed by health care improvement (Bero et al. 1998; Grol 2001). Several studies have identified barriers to guideline adherence (Cabana et al. 1999; Wensing, Grol, and Eccles 2005; Carlsen, Glenton, and Pope 2007; Lugtenberg, Burgers, and Westert 2009). These barriers can be classified

into *patient related barriers*, *physician related barriers*, *guideline related barriers*, and *organizational barriers* (Cabana et al. 1999). Implementation strategies tailored to barriers are known to be the most effective in improving guideline adherence (Grol 1997; Grimshaw et al. 2004). Hence, there is a need to gain insight into the application of the ESHRE guideline *Management of Women with Endometriosis* in daily practice (i.e. the *actual care*) and the potential barriers to guideline adherence. By connecting guideline evidence to daily practice, *quality indicators* are a suitable tool for measuring and monitoring the actual care and potential barriers (Grol, Baker, and Moss 2002). According to the ESHRE manual and additional literature (Mourad et al. 2007; Dancet et al. 2013; Luitjes et al. 2013), the first step in the development of *quality indicators* is to select a compact set of recommendations on which to focus, i.e. *key recommendations*. Experts strongly recommend to involve both professionals and patients in this selection procedure, since patients and professionals conceivably have different views regarding the best quality of care (Kotter et al. 2013; Krahn and Naglie 2008; Uphoff et al. 2012; den Breejen EM 2013). When developed, these key recommendations can be translated and validated into *quality indicators*. Measuring and monitoring the actual care can help to better implement the ESHRE Guideline in European hospitals and improve endometriosis care.

The aim of this study was 2-fold. First, the selection of a compact set of key recommendations as a first step in the development of quality indicators. Secondly, to detect differences in perspectives between patients and professionals regarding essential endometriosis care.

## Materials and methods

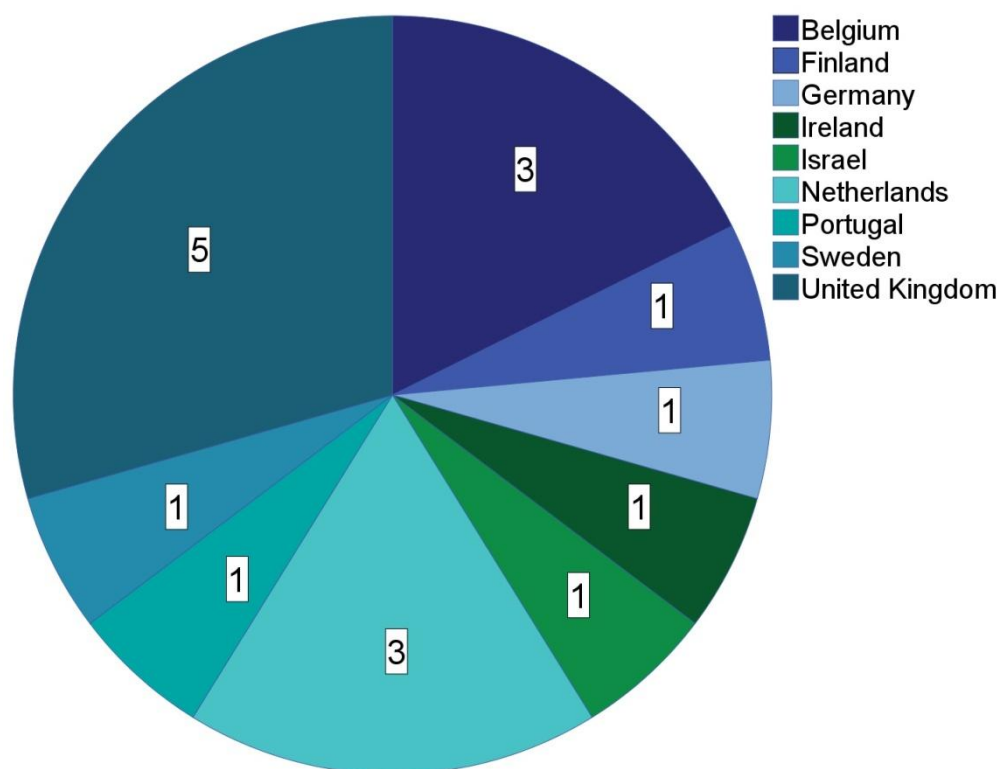
### Setting

A basic *RAND Delphi procedure* (Dalkey 1969; Fitch, Bernstein, and Aguilar 2001; Boulkedid et al. 2011) was used to develop a set of key recommendations, suitable for transcription into quality indicators, and based on the ESHRE Guideline *Management of Women with Endometriosis* (Dunselman et al. 2014). The Delphi procedure is considered an accepted methodology for the selection of key recommendations and development of quality indicators in health care (Campbell et al. 2003). In this systematic, stepwise method, evidence based information is combined with the individual opinion of

experts and aggregated into group consensus (Campbell et al. 2003; Boukdedid et al. 2011; Diamond et al. 2014; Kotter, Blozik, and Scherer 2012). In this study, two questionnaire rounds and one agreement round were performed to achieve panel consensus on the essential aspects of endometriosis care. Panel members were polled individually and anonymously. Opinions were swayed via repetitive feedback after each round, thus avoiding the negative social influences associated with face-to-face discussion (Fitch, Bernstein, and Aguilar 2001). Questionnaires were conducted with SurveyMonkey®. Possibilities to add comments on recommendations were provided in each questionnaire. Invitations and reminders were sent via SurveyMonkey®. All scores were listed in a database created with SPSS version 22.0. The consensus procedure took place between September 2015 and December 2015.

### **Composition of the expert panel**

To enhance the acceptance of the key recommendations in clinical practice, the expert panel consisted of a representative diversity of international patients and professionals. Patients and professionals were selected by their expertise and knowledge of endometriosis and ability to communicate in English. Eligible experts were medical doctors with a longstanding experience in the management of women with endometriosis and member of the *ESHRE Guideline Development Group*, and endometriosis patients with a prominent role in a patient organization. Thus, all European patient organizations (n=27) and medical doctors of the *ESHRE Guideline Development Group* (n=12) were informed and invited to join this study by E-mail. A total number of 10 patients and 11 professionals from 9 different European countries gave consent to participate in this study, forming an international expert panel of 21 members (**Figure I**).



**Figure I** The number of panel members per country.

## Selection of key recommendations

The selection of key recommendations consisted of six steps: (i) *extraction and classification of recommendations*, (ii) *first questionnaire round*, (iii) *data analysis of the first round*, (iv) *second questionnaire round*, (v) *data analysis of the second round*, and (vi) *approval of selected recommendations*. The steps taken in this Delphi method have been visualized in **Figure II** and described below.

### Step I - Extraction and classification of recommendations

Two authors (MS and NV) extracted 83 unique recommendations from the online version of ESHRE Guideline *Management of Women with Endometriosis*, published September 2013. Subsequently, two authors (MS and WN) distributed these recommendations into four domains, based on guideline chapters: (i) *diagnosis*, (ii) *treatment of endometriosis-associated pain*, (iii) *treatment of endometriosis-associated infertility* (including *treatment of endometriosis-associated infertility* and *medical assisted*

reproduction), and (iv) *miscellaneous topics* (including *menopause in women with endometriosis*, *asymptomatic endometriosis*, *prevention of endometriosis* and *endometriosis and cancer*).

## **Step 2 - First questionnaire round (Delphi round I)**

In the first round all 83 recommendations were presented to the expert panel in an online questionnaire. Panel members were asked to assess all recommendations individually on relevance. Relevance was graded by the experts in response to the following question; “*To what extent is the following guideline recommendation an important determinant to ensure or improve the quality of endometriosis care in European hospitals*” on a 9-point Likert scale ranging from 1 (extremely irrelevant) to 9 (extremely relevant). All participants had access to the ESHRE Guideline and the patient version of this guideline for supporting evidence or background information during the rating process. An example of the score sheet is provided in online **Appendix II**.

Secondly, the experts were asked to provide for each domain a top-3 ranking of recommendations they considered to be the most relevant in the contribution to high quality of endometriosis care, in order to promote discrimination between recommendations with a high Likert score.

## **Step 3 - Data analysis of the first round**

The results of the first round were analyzed using a-priori defined consensus criteria based on Campbell's criteria (Campbell 2000). These criteria include a median score of 8 or higher and panel agreement. Panel agreement was defined as in the case in which 75% or more of the individual scores was in the highest tertile of the scale (Likert score 7-9) and the other 25% of the scores was divided over the remaining two tertiles (Likert score 1-6). Previous studies (Mourad et al. 2007; van den Boogaard et al. 2010; Stienen et al. 2011; Uphoff et al. 2012; den Breejen EM 2013; Dancet et al. 2013; Luitjes et al. 2013; Woiski et al. 2015) have showed that using these two criteria only, often does not provide enough discrimination. Therefore, a third criterion was added: recommendations should have at least 20% of the maximum top-3 score. Points were awarded to each top-three ranking position, with number 1 position = 3 points, number 2 position = 2 points, and number 3



position = 1 point. These points were converted into percentages based on the maximum top 3 score. The maximum top 3 score was defined as the number of participants multiplied by the points awarded to a number one position. The study investigator (MS) combined the three criteria as described above and converted them into three possible outcomes: 'selected', 'rejected' and 'no consensus'. Recommendations that met all three criteria were classified as 'selected', those who met none of the criteria as 'rejected', and the remaining recommendations as 'no consensus'. The 'no consensus' recommendations were the input for the second questionnaire round. Finally, a Mann–Whitney *U* test was used to investigate potential different scoring behaviors between patients and professionals. An example of the consensus methodology is provided in online **Appendix III**.

#### **Step 4 - Second questionnaire round (Delphi round 2)**

The second round started with an overview of the 11 selected recommendations and the 42 'no consensus' recommendations. The spread of opinions in the 42 remaining 'no consensus' recommendations was visualized in box-and-whisker plots, showing the differences in median scores and overall scores between patients and professionals. We conducted for each panel member a personal questionnaire with their individual scores of the first round visualized in the box-and-whisker plots, encouraging them to revise their opinion in light of the replies of the other panel members. An example of a personalized box-and-whisker plot is provided in online **Appendix IV**. All panel members, including the non-responders of step 2, were once again asked to assess the 42 'no consensus' recommendations *on relevance on a 9-point Likert scale*. Secondly, the experts were asked for their approval of the 11 preselected key recommendations. Moreover, each panel member had the possibility to add one extra 'no consensus' recommendation per domain or to agree with the current selection.

#### **Step 5 - Data analysis of the second round**

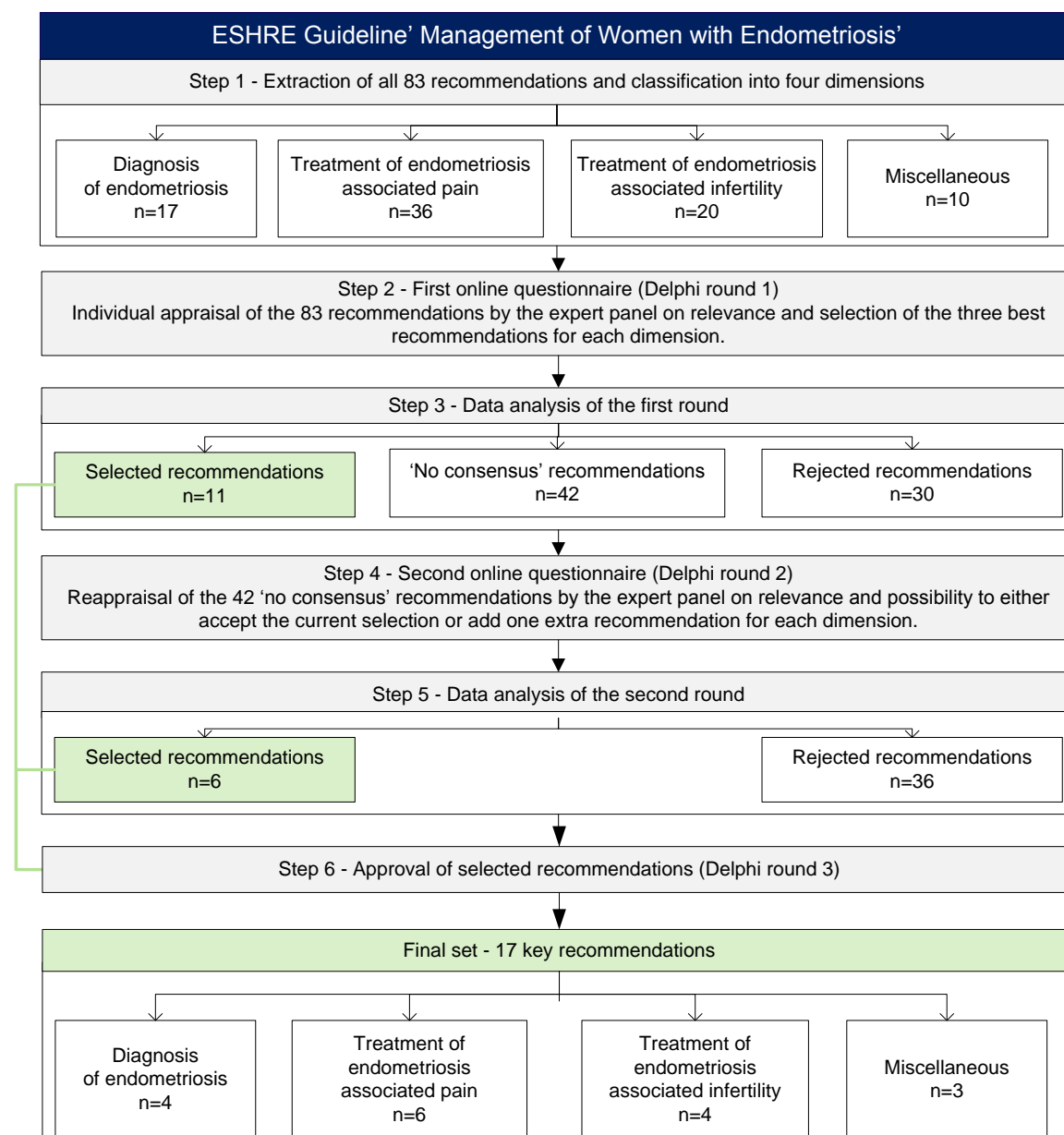
The 2 a-priori defined consensus criteria as defined in step 3 were used to analyze the results of the second questionnaire round. The selection of a recommendation by at least 2 experts was added as a

third criterion. These 3 criteria were combined and converted into 2 possible outcomes: 'selected' or 'rejected'. Recommendations that met all criteria were classified as 'selected' and the remaining recommendations as 'rejected'.

As in the first round, a Mann–Whitney *U* test was used to investigate differences in scoring behavior between patients and professionals.

#### **Step 6 – Approval of selected recommendations (Delphi round 3)**

Finally, all 21 panel members, including those who did not participate in the selection rounds, received an overview of the selected recommendations by E-mail. Experts were asked to approve the final set and were provided with a last opportunity to make remarks. Comments were discussed by the authors.



**Figure II** A step-wise basic RAND Delphi method was used to develop key recommendations for endometriosis care.

## Main outcome measures

The primary outcome measure was one set of key recommendations based on the ESHRE guideline and experts' opinion. To form the final set, the selected recommendations' of step 2 were supplemented with the selected recommendations of step 4 and approved by the expert panel in

step 6. The secondary outcome measure is the difference in perspective between patients and professionals according to the best quality of endometriosis care.

## **Ethical considerations**

In this study, formal ethical approval from a medical ethical committee was not required, since we did not use any patient medical data. Because all participants were adults, chosen on their expertise and willingness to participate, they were not considered vulnerable. Patients and professionals were informed by E-mail about the purpose and aim of this study, the procedures to be followed, the anticipated time commitment, and contact details for any questions. All participants gave consent before inclusion in this study. Withdrawal from the study was possible any time. One person (MS) collected all data to respect the privacy of the participants. Names of the participants were not linked to their responses in the questionnaire feedback. Responses were collected and analyzed anonymously. Safe storage of all data was provided.

## **Results**

### ***Composition of the expert panel***

The participating 7 patients represented 6 different countries: Belgium (n=2), Finland (n=1), Ireland (n=1), The Netherlands (n=1), Sweden (n=1), and United Kingdom (n=1). Most of the patients (57%) had been diagnosed with endometriosis more than 10 years ago. A majority (86%) experienced a delay in diagnosis of at least 5 years. All 7 patients had undergone surgery as a treatment, of them 86% received additional treatment for endometriosis-associated pain, and 43% received additional treatment for endometriosis-associated infertility.

The participating professionals also represented 6 different countries: Belgium (n=1), Germany (n=1), Israel (n=1) The Netherlands (n=3), Portugal (n=1), and United Kingdom (n=4). Most of the experts were gynaecologists specialized in endometriosis (91%) with a subspecialization in reproductive health (73%) or surgery (27%). One of the experts was a medical doctor and senior scientist

specialized in reproductive health and indicator development. Further characteristics of the panel members and descriptive data is provided in online **Appendix I**.

### **Step 1 - Extraction and classification of recommendations**

Altogether, 83 recommendations were extracted from the ESHRE Guideline and distributed into four domains: (i) *diagnosis* (n=17), (ii) *treatment of endometriosis-associated pain* (n=36), (iii) *treatment of endometriosis-associated infertility* (n=20), and (iv) *miscellaneous topics* (n=10).

### **Step 2 - First questionnaire (Delphi round 1)**

In the first Delphi round, 17 (81%) out of 21 experts who had given consent to participate, completed the first online questionnaire. The response rates were 70% (n=7) for patients and 91% (n=10) for professionals (**Figure III**). Reasons for not responding were time constraints among professionals and language related among patients. Non responders were sent a reminder after 2 and 3 weeks. Additional clarification about the aim of the study and time investment was provided to a professional. None of the participating patients needed additional clarification or information. One patient filled out the first questionnaire during the second round and could therefore not be included in the data analysis of the first round. The time needed for participants to fill out the first questionnaire was 0:33:59 hours for professionals (0:13:21 – 0:58:11 hours) and 0:56:29 hours for patients (0:22:28 – 1:52:43 hours).

### **Step 3 - Data analysis of the first round**

Recommendations were selected as potential key recommendations if they matched all the criteria described previously. Data analysis resulted in the selection of 11 recommendations (13%) (**Table I**), and rejection of 30 recommendations (36%). In 42 recommendations (51%) consensus could not be reached. The Mann-Whitney *U* test showed significant differences in the distribution of ratings between patients and professionals in 7 recommendations (**Table II**). All these 7 recommendations were rated higher by patients than professionals. By using the consensus criteria, 1 of these 7

recommendations was selected despite the inconsistency in ratings between patients and professionals, showing a prominent patient's interest in good communication.

#### **Step 4 - Second questionnaire (Delphi round 2)**

In the second Delphi round, 16 (76%) out of 21 experts, completed the second online questionnaire. The response rates were 60% (n=6) for patients and 91% (n=10) for professionals (**Figure III**). The expert panel consisted of the same experts as in the first round. Non responders were sent a reminder after 2 and 3 weeks. The time needed for participants to fill out the second questionnaire was 0:21:15 hours for professionals (0:10:26 – 0:45:33 hours) and 0:51:36 hours for patients (0:15:01 – 2:14:17 hours).

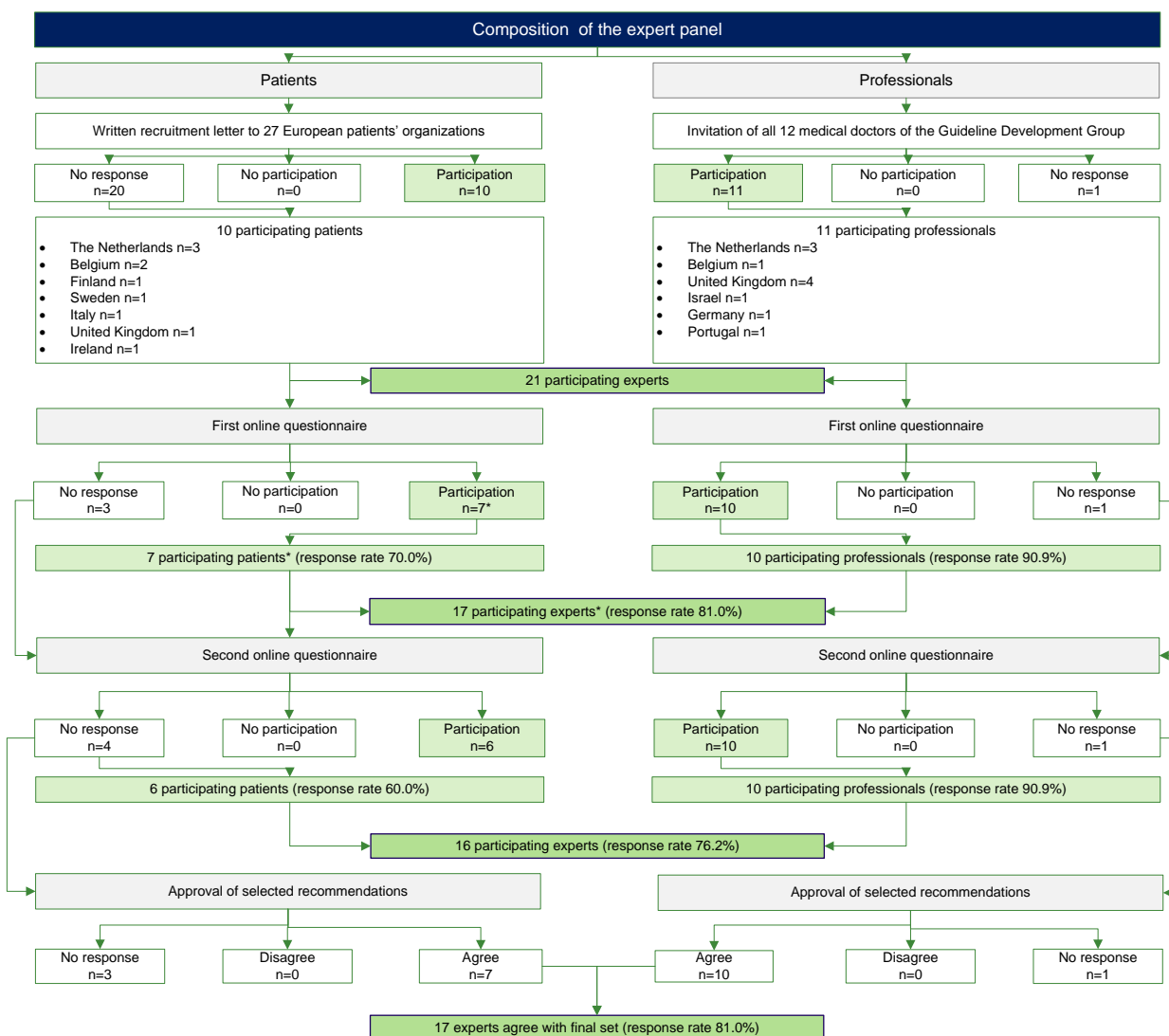
#### **Step 5 - Data analysis of the second round**

Recommendations were selected as potential key recommendations if they matched the criteria described previously. Data analysis resulted in the selection of 6 recommendations (14%), and rejection of remaining 30 recommendations (86%). The Mann-Whitney *U* test showed no difference in the distribution of ratings between patients and professionals, thus showing a shift towards consensus between both subgroups.

#### **Step 6 – Approval of selected recommendations (Delphi round 3)**

All 17 responding experts (81,0%) agreed with the final set of 17 selected recommendations for the management of women with endometriosis (**Table I**). Of all 17 key recommendations, 11 were selected in the first round and 6 were selected in the second round, representing 20% of the preliminary set of recommendations that was extracted from the ESHRE Guideline. The 17 key recommendations were equally divided over the 4 domains: (i) diagnosis of endometriosis (n=4/17, 24%), (ii) treatment of endometriosis-associated pain (n=6/36, 17%), treatment of endometriosis-associated infertility (n=4/20, 20%), and miscellaneous (n=3/10, 30%). Most key recommendations

(n=11) were *good practice points* (expert opinion), the remaining 6 were supported with level A evidence (meta-analysis, systematic review or multiple randomized controlled trials).



\* One of the patients filled out the first questionnaire during the second round and could therefore not be included in the data analysis of the first round.

**Figure III** Composition of the expert panel and participation of the panel members during the 3-round Delphi procedure.

**Table I** Final set of key recommendations for the management of women with endometriosis in European hospitals.

<i>Recommendations divided by chapter</i>	<i>Level of evidence</i>	<i>Selection round</i>
Clinicians should, in all women with endometriosis:		
<b>Diagnosis of endometriosis</b>		
I - Consider the diagnosis of endometriosis in the presence of gynecological symptoms such as: dysmenorrhea, non-cyclical pelvic pain, deep dyspareunia, infertility, fatigue in the presence of any of the above.	GPP	I
II - Consider the diagnosis of endometriosis in women of reproductive age with non-gynecological cyclical symptoms (dyschezia, dysuria, hematuria, rectal bleeding, shoulder pain).	GPP	I
III - Perform transvaginal sonography to diagnose or to exclude an ovarian endometrioma (Moore et al. 2002).	A	2
IV - Assess ureter, bladder, and bowel involvement by additional imaging if there is a suspicion based on history or physical examination of deep endometriosis, in preparation for further management.	GPP	2
<b>Treatment of endometriosis-associated pain</b>		
V - Counsel women with symptoms presumed to be due to endometriosis thoroughly, and empirically treat them with adequate analgesia, combined hormonal contraceptives or progestagens.	GPP	I
VI - Prescribe hormonal treatment [hormonal contraceptives (level B), progestagens (level A), antiprogestagens (level A), or GnRH agonists (level A)] as one of the options, as it reduces endometriosis-associated pain (Vercellini et al. 1993; Brown, Kives, and Akhtar 2012; Brown, Pan, and Hart 2010).	A-B	2
VII - Take patient preferences, side effects, efficacy, costs and availability into consideration when choosing hormonal treatment for endometriosis-associated pain.	GPP	I
VIII - Prescribe hormonal add-back therapy to coincide with the start of GnRH agonist therapy, to prevent bone loss and hypoestrogenic symptoms during treatment. This is not known to reduce the effect of treatment on pain relief (Bergqvist, Jacobson, and Harris 1997; Mäkräinen, Rönnerberg, and Kauppila 1996; Moghissi et al. 1998; Taskin et al. 1997).	A	2
IX - Surgically treat endometriosis when identified at laparoscopy, i.e. 'see and treat', as this is effective for reducing endometriosis-associated pain (Jacobson et al. 2009).	A	I
X - Refer women with suspected or diagnosed deep endometriosis to a centre of expertise that offers all available treatments in a multidisciplinary context.	GPP	I
<b>Treatment of endometriosis-associated infertility</b>		
XI - Perform operative laparoscopy (excision or ablation of the endometriosis lesions) including adhesiolysis, rather than performing diagnostic laparoscopy only in infertile women with AFS/ASRM stage I/II endometriosis, to increase ongoing pregnancy rates (Jacobson et al. 2010; Nowroozi et al. 1987).	A	I
XII - Perform excision of the endometrioma capsule, instead of drainage and electro coagulation of the endometrioma wall in infertile women with ovarian endometrioma undergoing surgery, to increase spontaneous pregnancy rates (Hart et al. 2008).	A	2
XIII - Counsel women with endometrioma regarding the risks of reduced ovarian function after surgery and the possible loss of the ovary. The decision to proceed with surgery should be considered carefully if the woman has had previous ovarian surgery	GPP	I
XIV - Use assisted reproductive technologies for infertility associated with endometriosis, especially if tubal function is compromised or if there is male factor infertility, and/or other treatments have failed.	GPP	2
<b>Miscellaneous topics</b>		
XV - Continue to treat women with a history of endometriosis after surgical menopause with combined estrogen/progestagen or tibolone, at least up to the age of natural menopause.	GPP	I
XVI - Fully inform and counsel women about any incidental finding of endometriosis.	GPP	I
XVII - Inform women with endometriosis, requesting information on their risk of developing cancer that 1) there is no evidence that endometriosis causes cancer, 2) there is no increase in overall incidence of cancer in women with endometriosis, and 3) some cancers (ovarian cancer and non-Hodgkin's lymphoma) are slightly more common in women with endometriosis.	GPP	I

ESHRE, European Society of Human Reproduction and Embryology; GPP, Good practice points; GnRH, Gonadotropin-releasing hormone; AFS/ASRM, American Fertility Society/American Society for Reproductive Medicine; Level A, Meta-analysis, systematic review or multiple randomized controlled trials (high quality); Level B, Meta-analysis, systematic review or multiple randomized controlled trials (moderate quality); Level C, Single randomized controlled trial, large non-randomized trial, case-control or cohort studies (moderate quality); Level D, Non-analytic studies, case reports or case series (high or moderate quality); GPP, Expert opinion.

**Table I** Final set of key recommendations for the management of women with endometriosis in European hospitals.



Recommendations divided by chapter

	Median score patients	Median score professionals	Mann-Whitney U	Conclusion
Clinicians should, in all women with endometriosis:				
<b>Diagnosis of endometriosis</b>				
Consider the diagnosis of endometriosis in the presence of gynecological symptoms such as: dysmenorrhea, non-cyclical pelvic pain, deep dyspareunia, infertility, fatigue in the presence of any of the above.	9	8	0.007*	No consensus
Be aware that the usefulness of 3D sonography to diagnose rectovaginal endometriosis is not well established.	9	5	0.007*	Rejected
Be aware that the usefulness of magnetic resonance imaging (MRI) to diagnose peritoneal endometriosis is not well established.	9	6	0.042*	Rejected
<b>Treatment of endometriosis-associated pain</b>				
Consider prescribing aromatase inhibitors in combination with oral contraceptive pills, progestagens, or GnRH analogues in women with pain from rectovaginal endometriosis refractory to other medical or surgical treatment, as they reduce endometriosis-associated pain.	8	5	0.013*	Rejected
Be aware that presacral neurectomy (PSN) is effective as an additional procedure to conservative surgery to reduce endometriosis-associated midline pain, but it requires a high degree of skill and is a potentially hazardous procedure.	8	7	0.040*	Rejected
Consider medical treatment of extragenital endometriosis when surgical treatment is difficult or impossible, to relieve symptoms.	8	7	0.040*	No consensus
<b>Miscellaneous</b>				
Fully inform and counsel women about any incidental finding of endometriosis.	9	7	0.008*	Selected

**Table II** Recommendations with a significant difference between patients' opinion and professionals' opinion in first questionnaire round.

## Discussion

As a first step in the development of quality indicators, an international expert panel of patients and medical professionals extracted 17 key recommendations from the ESHRE guideline *Management of Women with Endometriosis*. Differences between patients' perspective and professionals' perspective of essential endometriosis care were seen in the first Delphi round but disappeared in the following 2 rounds when opinions were swayed. The set of key recommendations covers all fields of endometriosis care, including diagnosis, treatment of endometriosis-associated pain, treatment of endometriosis-associated infertility and miscellaneous topics such as prevention, menopause and the relationship between endometriosis and cancer.

To our knowledge, this is the first study to select key recommendations as a first step in the development of quality indicators for the management of women with endometriosis. Remarkably, it

is one of the few studies where a combined panel of medical professionals as well as patients is involved in the selection procedure. It is well-known that patients have a invaluable merit when it comes to assessing the relevance or the weight of quality indicators (Pohontsch et al. 2015; Dancet et al. 2013; den Breejen EM 2013). Eventually, patients are the ultimate experts in patient-centeredness of care (Epstein and Street 2011; Grol 2001), which is one of core dimensions of *quality of care* (Institute of Medicine Committee on Quality of Health Care in 2001). The results of our study support the fact that there is a poor correlation between the initial patients' and professionals' perceptions regarding quality of care (Kotter et al. 2013; Krahn and Naglie 2008; Uphoff et al. 2012; den Breejen EM 2013; Aarts et al. 2011; van Empel et al. 2011). Hence, the set of key recommendations would have been different if only medical professionals were involved in the selection procedure. As an example, one of the key recommendations was selected by the patients despite a moderate popularity among the professionals, thereby overruling the opinion of the professionals. In this recommendation medical professionals are advised to fully inform and counsel women about any incidental finding of endometriosis. This is in line with previous studies reporting that professionals underestimate the importance of 'softer' dimensions of healthcare (e.g. respectful attitude and communication) and overestimate the importance of biomedical outcomes compared to patients (Laine et al. 1996; Rothwell et al. 1997; Mack et al. 2005; Wessels et al. 2010; van Empel et al. 2011). Furthermore, engagement of patients' perspective in healthcare not only leads to an higher patient-satisfaction, but is also proven to be effective clinically and economically (van Veenendaal, Voogdt-Pruis, and Raats 2015; Katz and Hawley 2013; Coulter 2008). Several studies on shared-decision making show correlations with improved health outcomes, better communication, and reductions in costs and unwarranted variations in care (Stacey et al. 2011; Oshima Lee and Emanuel 2013; Wennberg 2011).

The final set of 17 key recommendations fits the need for process and structural indicators besides outcome indicators (Nelen et al. 2007). Quality indicators can assess structures, processes and outcomes of healthcare (Donabedian 2005). Outcome indicators are mainly used by health insurances to judge hospitals in clinical practice. However, good outcome measures do not

necessarily stand for good quality of care, while the application of process and structural indicators in daily practice can lead to better outcome measures (Mant 2001; Smith 2008). Therefore, process and structural indicators are more valuable in healthcare improvement than outcome indicators, because they reveal the barriers in healthcare and provide clear pathways for action (Mainz 2004).

A key strength of this study is the combination of evidence and expert opinion, involving both patients and medical professionals from 9 different countries. A second strength is the number of selected key recommendations. Although previous studies show an average number of 20-50 quality indicators per condition (Boukdedid et al. 2011), ideally the development of quality indicators results in a compact set of  $\leq 5$  indicators per clinical area (van Doorn-Klomberg et al. 2013). We expect the set to be reduced by 10-20% during the practice test (Wollersheim et al. 2007). Other strengths of our study are the satisfying response rates and the panel size. According to the RAND manual, a panel size of at least 7-15 experts is recommended in order to develop a reliable set of indicators (Fitch, Bernstein, and Aguilar 2001). The number of participants in this study is comparable to panels used in other Delphi studies (Fitch, Bernstein, and Aguilar 2001; Boukdedid et al. 2011; Diamond et al. 2014; Kotter, Blozik, and Scherer 2012). Furthermore, our expert panel represented a robust sample of the most important stakeholders to make sure that all aspects of endometriosis care could be discussed. All professionals were medical doctors specialized in endometriosis and involved in the development of the ESHRE guideline. In our study fertile- and non-fertile patients took part, who had all undergone surgical treatment for endometriosis and at least one other therapy for endometriosis-associated pain or infertility. Literature shows that diversity of experts panel members leads to better performance as this may allow the consideration of different perspectives (Murphy et al. 1998). This diversity provides a suitable set of key recommendations for endometriosis care and should support a broad acceptance in daily practice internationally. A fourth strength is our methodologically strong Delphi design based on the RAND-manual, which is a renowned method for the development of quality indicators.

Although the response rates were good, some response bias may have occurred as not all panel members took part in all 3 rounds (Sica 2006) because of time constraints, complexity of the

Delphi procedure or language problems, experienced especially by patients. Therefore, the final set might reflect the opinion of the most motivated panel members and higher educated patients' representatives (Sica 2006). Furthermore, the homogeneity of the patients included in this study (e.g. all had a prominent role in a patient organization, all underwent surgery, and most of them had a long delay in diagnosis), may have influenced their scoring behavior. However, we consider the involved patients to be representative because of their diverse backgrounds and leadership in various international patient organizations (Kotter et al. 2013; Hermens et al. 2006; Ouwens et al. 2010; Kesmodel and Jolting 2011). Another discussion point is the level of evidence of the selected recommendations. Quality indicators are easier accepted in clinical practice if they are build on an high level of evidence (Mainz 2004). Our final set included 6 recommendations based on level A evidence. The remaining 11 recommendations were good practice points, formulated by members of the guideline development group. Although expert opinion is considered to be the lowest degree of evidence, the selection of these recommendations above recommendations with an higher level of evidence, shows their importance in daily practice (Eddy 1998). Finally, the feasibility of the selected key recommendations was not assessed in this study.

## **Implications for practice and future research**

In order to improve healthcare, quality indicators should be relevant, valid, reliable and feasible. All key recommendations in our final set are facially valid and reliable because they are based on an evidence-based guideline and selected on strict criteria by a considerable expert panel. The next step in the development of quality indicators should be to assess the feasibility of the key recommendations in clinical practice (Campbell et al. 2003). We suggest a pilot study in 1 or 2 hospitals to establish the applicability and measurability of our set of our key recommendations.

In future, the final set of quality indicators should be available for all hospitals to measure and monitor the actual endometriosis care and potential barriers to guideline adherence internationally. A multifaceted implementation strategy tailored to guideline-specific barriers should facilitate further improvement of endometriosis care (Grol and Grimshaw 2003). Future research has to establish

whether this implementation strategy has a positive influence on guideline adherence and healthcare outcomes.

## Conclusion

This study describes the systematic selection of key recommendations for endometriosis care by an international panel of patients and medical professionals as a first step in the development of quality indicators. The entire set of 17 key recommendations provides useful information on essential endometriosis care, regarding patients' and professional's perspective. In future, quality indicators can help to better implement the ESHRE guideline in hospitals by using strategies based on guideline-specific barriers, and thereby improve the quality of endometriosis care internationally. Furthermore, our results reinforce the importance of involving patients in the development of guidelines and quality indicators.

## Supplementary data

Supplementary data are available at <http://humrep.oxfordjournals.org>.

## Conflict of interest

The authors and members of the *EndoKey* study group declare that they have no conflict of interest in this procedure.

## Acknowledgements

We would like to acknowledge all patients for their participation in this selection procedure as well as the medical professionals as members of the *EndoKey* study group: Mrs. E.A. Andersson (Sweden), Mrs. C. Maasbach (The Netherlands), Mrs. M.P. Van de Ven – van Ham (The Netherlands), Mrs. M. Hollander (The Netherlands), Mrs. D. Paredes (Finland), Mrs. S. Pesci (Italy), Mrs. K. M. King (Ireland), Mrs. I. Banach (Belgium), Mrs. F. Jansen – van den Berg (Belgium), Dr G.A.J. Dunselman

(The Netherlands), Prof Dr C. M. Becker (United Kingdom), Prof Dr C. Calhaz-Jorge (Portugal), Prof Dr T. D'Hooghe (Belgium), Dr M.O. Heikinheimo (Finland), Prof Dr A.W. Horne (United Kingdom), Prof Dr L. Kiesel (Germany), Dr A. Nap (The Netherlands), Dr W.L.D.M. Nelen (The Netherlands), Dr A. Prentice (United Kingdom), Dr E. Saridogan (United Kingdom), and Dr D. Soriano (Israel).

### **Authors' roles**

M.S., W.N., G.D. and N.V. designed the research project. M.S. and N.V. composed the expert panel. M.S. conducted the surveys, led data collection, performed data analysis, and wrote this manuscript. The EndoKey members filled out the surveys. W.N., G.D. and N.V. contributed substantially to data interpretation and manuscript revisions. All authors read and approved the final manuscript.

### **Funding**

No funding has been received for the conduct of this study. Members of the EndoKey study group did not receive payment.

## References

- Aarts JW, Faber MJ, van Empel IW, Scheenjes E, Nelen WL, and Kremer JA. Professionals' perceptions of their patients' experiences with fertility care. *Hum Reprod* 2011: **26**; 1119-1127.
- Ballard K, Lowton K, and Wright J. What's the delay? A qualitative study of women's experiences of reaching a diagnosis of endometriosis. *Fertil Steril* 2006: **86**; 1296-1301.
- Bergqvist A, Jacobson J, and Harris S. A double-blind randomized study of the treatment of endometriosis with nafarelin or nafarelin plus norethisterone. *Gynecol Endocrinol* 1997: **11**; 187-194.
- Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, and Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. *BMJ* 1998: **317**; 465-468.
- Boulkedid R, Abdoul H, Loustau M, Sibony O, and Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 2011: **6**; e20476.
- Brown J, Kives S, and Akhtar M. Progestagens and anti-progestagens for pain associated with endometriosis. *Cochrane Database Syst Rev* 2012: **3**.
- Brown J, Pan A, and Hart RJ. Gonadotrophin-releasing hormone analogues for pain associated with endometriosis. *Cochrane Database Syst Rev* 2010: **CD008475**.
- Cabana MD, Rand CS, Powe NR, and et al. Why don't physicians follow clinical practice guidelines?: A framework for improvement. *JAMA* 1999: **282**; 1458-1465.
- Campbell SM, Braspenning J, Hutchinson A, and Marshall MN. Research methods used in developing and applying quality indicators in primary care. *BMJ* 2003: **326**; 816-819.
- Campbell SM, Cantrill, J. A., Roberts, D. Prescribing indicators for UK general practice: Delphi consultation study. *BMJ* 2000: **321**; 425-428.
- Carlsen B, Glenton C, and Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines. *Br J Gen Pract* 2007: **57**; 971-978.

- Coulter AP, S.; Askham, J.;. Where are the patients in decision-making about their own care? . *Policy brief, written for the WHO European Ministerial Conference on Health Systems. Estonia.* 2008.
- Culley L, Law C, Hudson N, Denny E, Mitchell H, Baumgarten M, and Raine-Fenning N. The social and psychological impact of endometriosis on women's lives: a critical narrative review. *Human Reproduction Update* 2013.
- Dalkey NCB, B.; Cochran, S. The Delphi Method III: Use of self ratings to improve group estimates. *Santa Monica, CA: Rand.* 1969.
- Dancet EA, D'Hooghe TM, Spiessens C, Sermeus W, De Neubourg D, Karel N, Kremer JA, and Nelen WL. Quality indicators for all dimensions of infertility care quality: consensus between professionals and patients. *Hum Reprod* 2013; **28**; 1584-1597.
- de Graaff AA, D'Hooghe TM, Dunselman GA, Dirksen CD, Hummelshoj L, Consortium WE, and Simoens S. The significant effect of endometriosis on physical, mental and social wellbeing: results from an international cross-sectional survey. *Hum Reprod* 2013; **28**; 2677-2685.
- den Breejen EM NW, Schol SF, et al. Development of guideline-based indicators for patient-centredness in fertility care: what patients add. *Hum Reprod* 2013; **28**; 987-996.
- Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, and Wales PW. Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology* 2014; **67**; 401-409.
- Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q* 2005; **83**; 691-729.
- Dunselman GA, Vermeulen N, Becker C, Calhaz-Jorge C, D'Hooghe T, De Bie B, Heikinheimo O, Horne AW, Kiesel L, Nap A, et al. ESHRE guideline: management of women with endometriosis. *Hum Reprod* 2014; **29**; 400-412.
- Eddy DM. Performance measurement: problems and solutions. *Health Aff (Millwood)* 1998; **17**; 7-25.
- Epstein RM and Street RL, Jr. The values and value of patient-centered care. *Ann Fam Med* 2011; **9**; 100-103.
- Eskenazi B and Warner ML. Epidemiology of endometriosis. *Obstet Gynecol Clin North Am* 1997; **24**; 235-258.



- 518 Fitch K, Bernstein SJ, and Aguilar MD. The RAND/UCLA Appropriateness Method User's Manual.  
519 2001.
- 520 Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP,  
521 Matowe L, Shirran L, et al. Effectiveness and efficiency of guideline dissemination and  
522 implementation strategies. *Health Technol Assess* 2004; **8**; iii-iv, 1-72.
- 523 Grol R. Personal paper: Beliefs and evidence in changing clinical practice 1997.
- 524 Grol R. Improving the quality of medical care: Building bridges among professional pride, payer profit,  
525 and patient satisfaction. *JAMA* 2001; **286**; 2578-2585.
- 526 Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice.  
527 *Med Care* 2001; **39**; 1i46-54.
- 528 Grol R, Baker R, and Moss F. Quality improvement research: understanding the science of change in  
529 health care. *Qual Saf Health Care* 2002; **11**; 110-111.
- 530 Grol R and Grimshaw J. From best evidence to best practice: effective implementation of change in  
531 patients' care. *The Lancet* 2003; **362**; 1225-1230.
- 532 Hart RJ, Hickey M, Maouris P, and Buckett W. Excisional surgery versus ablative surgery for ovarian  
533 endometriomata. *Cochrane Database Syst Rev* 2008; CD004992.
- 534 Hermens RP, Ouwens MM, Vonk-Okhuijsen SY, van der Wel Y, Tjan-Heijnen VC, van den Broek LD,  
535 Ho VK, Janssen-Heijnen ML, Groen HJ, Grol RP, et al. Development of quality indicators for  
536 diagnosis and treatment of patients with non-small cell lung cancer: a first step toward  
537 implementing a multidisciplinary, evidence-based guideline. *Lung Cancer* 2006; **54**; 117-124.
- 538 Institute of Medicine Committee on Quality of Health Care in A. Crossing the Quality Chasm: A  
539 New Health System for the 21st Century. 2001. National Academies Press (US), Washington  
540 (DC).
- 541 Jacobson TZ, Duffy JM, Barlow D, Farquhar C, Koninckx PR, and Olive D. Laparoscopic surgery for  
542 subfertility associated with endometriosis. *Cochrane Database Syst Rev* 2010; CD001398.
- 543 Jacobson TZ, Duffy JM, Barlow D, Koninckx PR, and Garry R. Laparoscopic surgery for pelvic pain  
544 associated with endometriosis. *Cochrane Database Syst Rev* 2009; CD001300.

- 545 Johnson NP and Hummelshoj L. Consensus on current management of endometriosis. *Hum Reprod*  
546 2013; **28**; 1552-1568.
- 547 Katz SJ and Hawley S. The value of sharing treatment decision making with patients: expecting too  
548 much? *JAMA* 2013; **310**; 1559-1560.
- 549 Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, Hummelshoj L, Prentice A,  
550 Saridogan E, Endometriosis ESIGf, et al. ESHRE guideline for the diagnosis and treatment of  
551 endometriosis. *Hum Reprod* 2005; **20**; 2698-2704.
- 552 Kesmodel US and Jolving LR. Measuring and improving quality in obstetrics--the implementation of  
553 national indicators in Denmark. *Acta Obstet Gynecol Scand* 2011; **90**; 295-304.
- 554 Klein S, D'Hooghe T, Meuleman C, Dirksen C, Dunselman G, and Simoens S. What is the societal  
555 burden of endometriosis-associated symptoms? a prospective Belgian study. *Reprod Biomed*  
556 *Online* 2014; **28**; 116-124.
- 557 Kotter T, Blozik E, and Scherer M. Methods for the guideline-based development of quality  
558 indicators--a systematic review. *Implement Sci* 2012; **7**; 21.
- 559 Kotter T, Schaefer FA, Scherer M, and Blozik E. Involving patients in quality indicator development -  
560 a systematic review. *Patient Prefer Adherence* 2013; **7**; 259-268.
- 561 Krahn M and Naglie G. The next step in guideline development: incorporating patient preferences.  
562 *JAMA* 2008; **300**; 436-438.
- 563 Laine C, Davidoff F, Lewis CE, Nelson EC, Nelson E, Kessler RC, and Delbanco TL. Important  
564 elements of outpatient care: a comparison of patients' and physicians' opinions. *Ann Intern*  
565 *Med* 1996; **125**; 640-645.
- 566 Lugtenberg M, Burgers JS, and Westert GP. Effects of evidence-based clinical practice guidelines on  
567 quality of care: a systematic review. *Qual Saf Health Care* 2009; **18**; 385-392.
- 568 Luitjes SH, Wouters MG, Franx A, Bolte AC, de Groot CJ, van Tulder MW, and Hermens RP.  
569 Guideline-based development of quality indicators for hypertensive diseases in pregnancy.  
570 *Hypertens Pregnancy* 2013; **32**; 20-31.

- 571 Mack JW, Hilden JM, Watterson J, Moore C, Turner B, Grier HE, Weeks JC, and Wolfe J. Parent and  
572 physician perspectives on quality of care at the end of life in children with cancer. *J Clin Oncol*  
573 2005; **23**; 9155-9161.
- 574 Mainz J. Quality indicators: essential for quality improvement. *Int J Qual Health Care* 2004; **16**; Suppl 1:  
575 i1-i2.
- 576 Mäkräinen L, Rönnerberg L, and Kauppila A. Medroxyprogesterone acetate supplementation  
577 diminishes the hypoestrogenic side effects of gonadotropin-releasing hormone agonist  
578 without changing its efficacy in endometriosis. *Fertil Steril* 1996; **65**; 29-34.
- 579 Mant J. Process versus outcome indicators in the assessment of quality of health care. *Int J Qual*  
580 *Health Care* 2001; **13**; 475-480.
- 581 Meuleman C, Vandenabeele B, Fieuws S, Spiessens C, Timmerman D, and D'Hooghe T. High  
582 prevalence of endometriosis in infertile women with normal ovulation and normospermic  
583 partners. *Fertil Steril* 2009; **92**; 68-74.
- 584 Moghissi KS, Schlaff WD, Olive DL, Skinner MA, and Yin H. Goserelin acetate (Zoladex) with or  
585 without hormone replacement therapy for the treatment of endometriosis. *Fertil Steril* 1998;  
586 **69**; 1056-1062.
- 587 Moore J, Copley S, Morris J, Lindsell D, Golding S, and Kennedy S. A systematic review of the  
588 accuracy of ultrasound in the diagnosis of endometriosis. *Ultrasound Obstet Gynecol* 2002; **20**;  
589 630-634.
- 590 Moradi M, Parker M, Sneddon A, Lopez V, and Ellwood D. Impact of endometriosis on women's  
591 lives: a qualitative study. *BMC Womens Health* 2014; **14**; 123.
- 592 Mourad SM, Hermens RP, Nelen WL, Braat DD, Grol RP, and Kremer JA. Guideline-based  
593 development of quality indicators for subfertility care. *Hum Reprod* 2007; **22**; 2665-2672.
- 594 Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, and Marteau T.  
595 Consensus development methods, and their use in clinical guideline development. *Health*  
596 *Technol Assess* 1998; **2**; i-iv, 1-88.

- 597 Nelen WL, Hermens RP, Mourad SM, Haagen EC, Grol RP, and Kremer JA. Monitoring reproductive  
598 health in Europe: what are the best indicators of reproductive health? A need for evidence-  
599 based quality indicators of reproductive health care. *Hum Reprod* 2007; **22**; 916-918.
- 600 Nnoaham KE, Hummelshoj L, Webster P, d'Hooghe T, de Cicco Nardone F, de Cicco Nardone C,  
601 Jenkinson C, Kennedy SH, and Zondervan KT. Impact of endometriosis on quality of life and  
602 work productivity: a multicenter study across ten countries. *Fertil Steril* 2011; **96**; 366-  
603 373.e368.
- 604 Nowroozi K, Chase JS, Check JH, and Wu CH. The importance of laparoscopic coagulation of mild  
605 endometriosis in infertile women. *Int J Fertil* 1987; **32**; 442-444.
- 606 Oshima Lee E and Emanuel EJ. Shared decision making to improve care and reduce costs. *N Engl J*  
607 *Med* 2013; **368**; 6-8.
- 608 Ouwens M, Hermens R, Hulscher M, Vonk-Okhuijsen S, Tjan-Heijnen V, Termeer R, Marres H,  
609 Wollersheim H, and Grol R. Development of indicators for patient-centred cancer care.  
610 *Support Care Cancer* 2010; **18**; 121-130.
- 611 Pohontsch NJ, Herzberg H, Joos S, Welte F, Scherer M, and Blozik E. The professional perspective on  
612 patient involvement in the development of quality indicators: a qualitative analysis using the  
613 example of chronic heart failure in the German health care setting. *Patient Prefer Adherence*  
614 2015; **9**; 151-159.
- 615 Rothwell PM, McDowell Z, Wong CK, and Dorman PJ. Doctors and patients don't agree: cross  
616 sectional study of patients' and doctors' perceptions and assessments of disability in multiple  
617 sclerosis. *BMJ* 1997; **314**; 1580-1583.
- 618 Sica GT. Bias in research studies. *Radiology* 2006; **238**; 780-789.
- 619 Simoens S, Dunselman G, Dirksen C, Hummelshoj L, Bokor A, Brandes I, Brodsky V, Canis M,  
620 Colombo GL, DeLeire T, et al. The burden of endometriosis: costs and quality of life of  
621 women with endometriosis and treated in referral centres. *Hum Reprod* 2012; **27**; 1292-  
622 1299.

- 623 Smith PCM, E.; Papanicolas, I.;. Performance measurement for health system improvement:  
624 experiences, challenges and prospects. 2008. WHO conference 'Health Systems. Health and  
625 Wealth', Estonia.
- 626 Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, Llewellyn-Thomas H, Lyddiatt  
627 A, Legare F, and Thomson R. Decision aids for people facing health treatment or screening  
628 decisions. *Cochrane Database Syst Rev* 2011; Cd001431.
- 629 Stienen JJ, Tabbers MM, Benninga MA, Harmsen M, and Ouwens MM. Development of quality  
630 indicators based on a multidisciplinary, evidence-based guideline on pediatric constipation.  
631 *Eur J Pediatr* 2011; **170**; 1513-1519.
- 632 Taskin O, Yalcinoglu AI, Kucuk S, Uryan I, Buhur A, and Burak F. Effectiveness of tibolone on  
633 hypoestrogenic symptoms induced by goserelin treatment in patients with endometriosis.  
634 *Fertil Steril* 1997; **67**; 40-45.
- 635 Uphoff EP, Wennekes L, Punt CJ, Grol RP, Wollersheim HC, Hermens RP, and Ottevanger PB.  
636 Development of generic quality indicators for patient-centered cancer care by using a RAND  
637 modified Delphi method. *Cancer Nurs* 2012; **35**; 29-37.
- 638 van den Boogaard E, Goddijn M, Leschot NJ, Veen F, Kremer JA, and Hermens RP. Development of  
639 guideline-based quality indicators for recurrent miscarriage. *Reprod Biomed Online* 2010; **20**;  
640 267-273.
- 641 van Doorn-Klomberg ALM, Braspenning JCCP, Feskens RCWP, Bouma MMDP, Campbell  
642 SMBAMAP, and Reeves DBP. Precision of Individual and Composite Performance Scores:  
643 The Ideal Number of Indicators in an Indicator Set. *Medical Care* 2013; **51**; 115-121.
- 644 van Empel IW, Dancet EA, Koolman XH, Nelen WL, Stolk EA, Sermeus W, D'Hooghe TM, and  
645 Kremer JA. Physicians underestimate the importance of patient-centredness to patients: a  
646 discrete choice experiment in fertility care. *Hum Reprod* 2011; **26**; 584-593.
- 647 van Veenendaal H, Voogdt-Pruis H, and Raats I. Samen beslissen is beter. *Tijdschrift voor*  
648 *praktijkondersteuning* 2015; **10**; 90-95.

- 649 Vercellini P, Trespidi L, Colombo A, Vendola N, Marchini M, and Crosignani PG. A gonadotropin-  
650 releasing hormone agonist versus a low-dose oral contraceptive for pelvic pain associated  
651 with endometriosis. *Fertil Steril* 1993; 75–79.
- 652 von Ferber L, Koster I, and Hauner H. [Costs of the treatment of hyperglycemia in patients with  
653 diabetes mellitus. The impact of age, type of therapy and complications: results of the  
654 German CoDiM study]. *Med Klin (Munich)* 2006; 101; 384-393.
- 655 Wennberg JE. Time to tackle unwarranted variations in practice. *BMJ* 2011; 342.
- 656 Wensing M, Grol R, and Eccles MP. Improving patient care. The implementation of change in clinical  
657 practice 2005. Butterworth-Heinemann
- 658 Wessels H, de Graeff A, Wynia K, de Heus M, Kruitwagen CL, Teunissen SC, and Voest EE. Are  
659 health care professionals able to judge cancer patients' health care preferences correctly? A  
660 cross-sectional study. *BMC Health Serv Res* 2010; 10; 198.
- 661 Woiski MD, Scheepers HC, Liefers J, Lance M, Middeldorp JM, Lotgering FK, Grol RP, and Hermens  
662 RP. Guideline-based development of quality indicators for prevention and management of  
663 postpartum hemorrhage. *Acta Obstet Gynecol Scand* 2015; 94; 1118-1127.
- 664 Wollersheim H, Hermens R, Hulscher M, Braspenning J, Ouwens M, Schouten J, Marres H, Dijkstra  
665 R, and Grol R. Clinical indicators: development and applications. *Neth J Med* 2007; 65; 15-22.
- 666
- 667

**Appendix I** Baseline characteristics of panel members.

Characteristics		Value, n (%)	
Gender	Male	9 (82%)	669
	Female	2 (18%)	
Age	35-44 years	2 (18%)	672
	45-54 years	3 (27%)	
	55-64 years	6 (55%)	
Country	Belgium	1 (9%)	675
	Germany	1 (9%)	
	Israel	1 (9%)	
	The Netherlands	3 (27%)	
	Portugal	1 (9%)	
	United Kingdom	4 (36%)	
Specialization	Gynaecology	10 (91%)	681
	Other	1 <sup>a</sup> (9%)	
Sub specialization	Reproductive Health	9 (82%)	682
	Surgery	2 (18%)	
Years of working experience	<10	1 (9%)	684
	10 -19 years	6 (55%)	
	20 – 29 years	2 (18%)	
	≥ 30 years	2 (18%)	

<sup>a</sup> Senior scientist reproductive health.**Appendix II** Example of the score sheet.

**To what extent is the following guideline recommendation an important determinant to measure the quality of endometriosis care?**

*Clinicians should consider the diagnosis of endometriosis in the presence of gynecological symptoms such as: dysmenorrhea, non-cyclical pelvic pain, deep dyspareunia, infertility, fatigue in the presence of any of the above.*

1                      2                      3                      4                      5                      6                      7                      8                      9  
extremely                      neutral                      extremely  
irrelevant

☐                      ☐                      ☐                      ☐                      ☐                      ☐                      ☐                      ☐

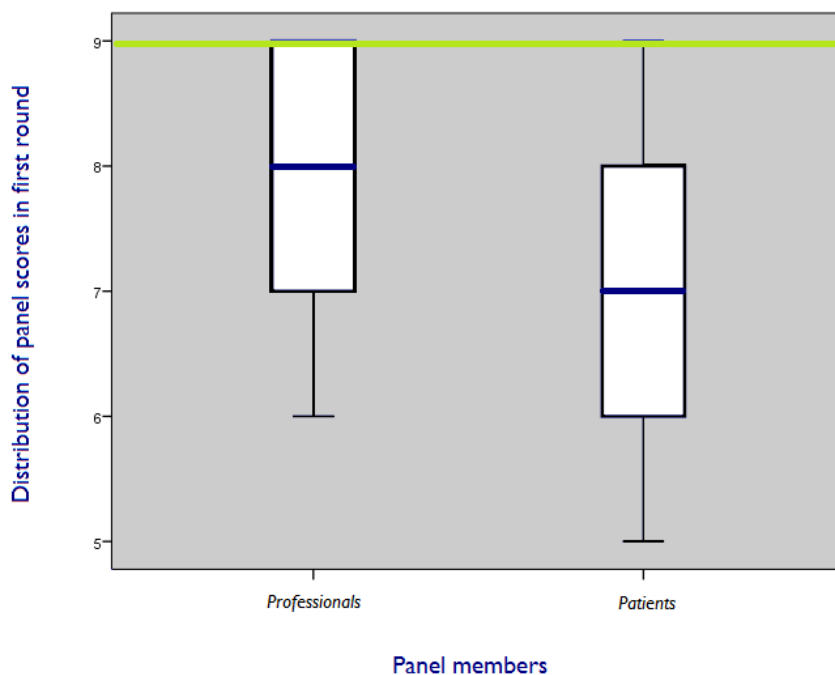
Comments:

### Appendix III Example of the consensus methodology.

Example of recommendation	Relevance score on 9-point Likert scale									Tertile score			Top 3 score (% of max)	Median	Conclusion
	1	2	3	4	5	6	7	8	9	1st	2nd	3rd			
Consider the diagnosis of endometriosis in the presence of gynaecological symptoms such as: dysmenorrhea, non-cyclical pelvic pain, deep dyspareunia, infertility, fatigue in the presence of any of the above.	0	0	0	0	1	0	2	1	12	0,0 %	6,3 %	93,8 %	68,8%	9.0	Selected
Consider the diagnosis of ovarian endometrioma in women with adnexal masses detected during clinical examination.	0	0	0	0	2	6	1	4	3	0,0 %	50,0 %	50,0 %	0,0%	6.5	Rejected
Perform transvaginal sonography to diagnose or to exclude an ovarian endometrioma.	0	0	0	0	3	1	1	4	7	0,0 %	25,0 %	75,0 %	6,3%	8.0	No consensus
Perform a laparoscopy to diagnose endometriosis, although evidence is lacking that a positive laparoscopy without histology proves the presence of endometriosis.	0	0	1	0	1	3	2	3	6	6,3 %	25,0 %	68,8 %	18,8%	8.0	No consensus

### Appendix IV Example of a personalized box-and-whisker plot.

Example of a box-and-whisker plot used in the second questionnaire to visualize a panel member's opinion in the light of the replies of the other panel members and the differences in ratings between patients and professionals.



The box includes 50% of the panel scores. The whiskers extend to the minimum and maximum panel scores. The thick blue line represents the median score of patients and professionals. The green line represents the panel member's score in the first round.