

# Pain scoring in endometriosis: entry criteria and outcome measures for clinical trials. Report from the Art and Science of Endometriosis meeting

Katy Vincent, M.B.B.S.,<sup>a</sup> Stephen Kennedy, M.D.,<sup>a</sup> and Pamela Stratton, M.D.<sup>b</sup>

<sup>a</sup> Nuffield Department of Obstetrics and Gynaecology, John Radcliffe Hospital, Headley Way, Oxford, United Kingdom; and

<sup>b</sup> Reproductive Biology and Medicine Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland

Standardized entry criteria and outcome measures for clinical trials in endometriosis-related pain would facilitate the comparison of trial results and the production of systematic reviews, improving evidence-based practice in this area. This report summarizes the recommendations from an international meeting for these criteria. (Fertil Steril® 2008; ■: ■–■. ©2008 by American Society for Reproductive Medicine.)

**Key Words:** Endometriosis, pain, quality of life, dysmenorrhea

An international meeting was convened by the National Institutes of Health (NIH), in collaboration with the American Society for Reproductive Medicine (ASRM), with the aim of establishing entry criteria and outcome measures for use in international clinical trials in endometriosis with regard to pain symptoms. A panel of invited scientists and clinicians from the United Kingdom, United States, and Italy gave presentations. Responses to 13 key questions that were raised by the moderators were then sought from the panel and the audience. A document summarizing the meeting recommendations was circulated to attendees and speakers for comment. The revised consensus document was approved by professional bodies such as the Special Interest Group on Endometriosis of the ASRM and European Society for Human Reproduction and Embryology.

## OVERVIEW OF THE PRESENTATIONS

All agreed that standardized entry and outcome criteria would be of benefit to clinicians, academics, industry, and patients to allow for comparisons across trials and different treatments. It was readily apparent that "... perhaps the most common error committed by clinical researchers is to dismiss existing scales too lightly, and embark on the development of a new instrument with an unjustifiably optimistic and naïve expectation that they can do better" (1).

The Biberoglu and Behrman (B&B) scale (2) was presented by David Olive. Although widely used in clinical stud-

Received March 4, 2008; revised and accepted September 17, 2008.

K.V. has received funding from Pfizer. S.K. has received grant support from Pfizer and is on the advisory boards of TAP Pharmaceuticals, Takeda, and Bayer Schering-Plough. P.S. has nothing to disclose.

This study was supported by the Intramural Program of the National Institute of Child Health and Human Development and unrestricted educational grants from Ferring, Neurocrine, Pfizer, Schering, Takeda, and TAP Pharmaceuticals.

Reprint requests: Dr. Katy Vincent, Nuffield Department of Obstetrics and Gynaecology, John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU (FAX: 00-44-1865-222727; E-mail: [kvincent@fmrib.ox.ac.uk](mailto:kvincent@fmrib.ox.ac.uk)).

ies, it has considerable limitations. The B&B asks questions about function and quality of life and thus is not a pain scale. It includes three symptoms, dysmenorrhea, dyspareunia, and chronic pelvic pain, and two signs, pelvic tenderness and induration, each of which are graded on a scale from 0 to 3 (or 4), with higher numbers indicating more severe symptoms. However, the score is inaccurately inflated when the highest score of 4 is given to women without symptoms such as those with amenorrhea or those who are not sexually active.

There is also no consistency as to whether it will be completed by the patient herself, administered by the physicians and study staff who collect the patient's information, or collected as the clinician's impression of the patient's symptoms. Also lacking is a standard for the symptoms that will include women in studies or indicate that a treatment has succeeded.

Studies have inconsistently reported the results of the scale, with some reporting only the individual scores for the symptoms and signs (3), some reporting only the sum of the symptoms (4) and others reporting the sum of all the signs and symptoms (5). In summary, it appears not to have been used or administered in a consistent manner, and it has never been validated or shown to be reproducible.

Paolo Vercellini presented information regarding the association between lesions and pain symptoms as background in determining both entry criteria and outcome measures. In his published metaanalysis of 1054 consecutive patients undergoing first-line conservative or definitive surgery in his unit between 1996 and 2002 (6), he concluded that neither the ASRM classification stages nor endometriomas are associated with pain severity and that the ASRM classification stage is not predictive of postoperative results or symptom recurrence. This is in agreement with data from others around the world, such as Szendei and colleagues in Hungary (7) and Mahmood et al. in Scotland (8). Instead, pain appears to be associated with deeply infiltrating lesions, the distance

between nerve fibers and implants, and the number of nerve fibers within lesions (9).

The neuroscience of pain and endometriosis was presented by Karen Berkley using her translational research of a rat animal model (10). She illustrated that endometriosis and its associated pelvic visceral and muscle pain often occur alongside other painful conditions in widely disparate body regions. Her animal and human data suggest that the ectopic growths are innervated and may produce algogenic agents peripherally that contribute to engaging the central nervous system (CNS) in generating pain symptoms. Importantly, she demonstrated that the experience of pain is a CNS phenomenon that arises from the intercommunication or matrix of connections in the brain. Since the CNS has a great deal of plasticity, the pain mechanisms related to endometriosis likely include central hormonal modulation, central sensitization, and remote central sensitization. These mechanisms are not unique to endometriosis and likely apply to other types of pain.

An overview of the different questionnaires used to assess quality of life (QoL) related to endometriosis was given by Crispin Jenkinson. While many studies have shown that the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) has high internal consistency and is useful in other medical conditions, there has not been a good correlation between QoL and pain intensity and/or use of medications in women with endometriosis (11–13). Similarly, the Euro-QOL EQ-5D has been used but has not been validated in women with endometriosis (12). The only QoL scale that has been validated for use in women with endometriosis is the Endometriosis Health Profile 30 (EHP-30) (14–16).

Charles Cleeland presented how to assess endometriosis pain from the perspective of a clinical expert in pain. To treat any type of pain, the following should be established: the sever-

ity, quality, location, and temporal pattern of the pain, as well as how it interferes with activities; its response to prior treatments; the adverse effects of treatment; and whether the pain is somatic, neuropathic, or visceral. Since pain is a subjective state, some have wondered whether they can trust their patients' ratings. Studies suggest that patient ratings are reliable.

Dr. Cleeland also presented different types of pain scales. The verbal rating scale or visual analogue scale is a scale in which 0 corresponds to no symptoms and 10 to the worst pain imaginable. The strengths of scales are that they are easy to administer and score, are sensitive to treatment effects, and correlate with other intensity measures. Their weaknesses include that they have limited response categories, assume equal intervals between adjectives, and are not appropriate for low literacy patients. The Brief Pain Inventory, which comprises 11 questions and a pain drawing (17), is quick to complete, uses 0–10 scales so it is easy to complete, measures both pain severity and interference, and is very sensitive to effective treatment.

Dr. Cleeland presented the Initiative on Methods, Measurement and Pain Assessment in Clinical Trials (IMMPACT) recommendations for clinical outcomes in pain trials (Table 1), which include pain measured in 0–10 scales, physical functioning (Brief Pain Inventory), emotional functioning (Beck, POMS), participant ratings of improvement or satisfaction (7- or 9-point scale), symptoms and adverse events, patient disposition (who drops out and why), and responder analysis (i.e., those with mild pain or 30% reduction) (18). He advised that clinical studies should carefully define responders, use 0–10 scales, pick time points for measurement that are relevant to the condition under study, use latency of pain relief as a secondary outcome, and use pain symptoms as eligibility criteria.

**TABLE 1****Summary of IMMPACT recommendations (adapted from 22).**

Core outcome measures	Tool
Pain	11-point (0–10) NRS of pain intensity; use of rescue analgesics; categorical rating of pain intensity in circumstances in which numerical ratings may be problematic
Physical functioning (either measure)	Multidimensional Pain Inventory Interference Scale; Brief Pain Inventory interference items
Emotional Functioning (at least one measure)	Beck Depression Inventory; Profile of Mood States
Participant ratings of global improvement and satisfaction with treatment	Patient Global Impression of Change
Symptoms and adverse events	Passive capture of spontaneously reported adverse events and symptoms and use of open-ended prompts
Participant disposition	Detailed information regarding participant recruitment and progress through the trial, including all information specified in the CONSORT guidelines

*Vincent. Pain outcome in endometriosis studies. Fertil Steril 2008.*

Data analysis and interpretation were presented by Charles Cao and stressed the importance of collecting and analyzing data systematically. Mary Lou Ballweg presented the Endometriosis Association's new physician's DVD portraying the experience of endometriosis from the patient point of view as well as eight key points indicating what is important to patients. Patients want [1] to be believed (tired of being told pain is in their heads or normal), [2] the physician to take the time to hear what the problem is before interrupting, [3] acknowledgement/validation, [4] an explanation for the problem, [5] their pain addressed! [6] a treatment and pain management plan addressing their key concerns, [7] a focus on functionality rather than just lesions, and [8] assistance and referrals to other helpful specialties.

There was consensus that the IMMPACT recommendations should be applied to endometriosis. There followed a debate about QoL measures, which may be valuable in a disease such as endometriosis, even though they have been shown not to be useful in cancer trials. Also debated were how to interpret QoL data and what constituted a clinically meaningful effect (particularly whether it should be defined by patients and/or doctors).

The following 13 questions were then discussed in the light of the presentations.

## THE 13 QUESTIONS

### 1. How do we define "endometriosis" for clinical trials?

There were marked differences of opinion between clinicians and scientists and between participants from different parts of the world that were unresolved. The definitions suggested were

- a. Clinical picture: symptoms suggestive of endometriosis
- b. Surgical diagnosis: endometriosis seen at laparoscopy/laparotomy
- c. Histological diagnosis
- d. Response to treatment: pain relief with GnRH agonist

It was considered necessary for clinical trials in endometriosis to include only patients with a surgical diagnosis rather than those with a presumed diagnosis on the basis of either symptoms or a response to GnRH agonist treatment. Although histological confirmation is desirable (19), it was decided not to include this as a requirement for entry into trials. This is consistent with both U.S. (20) and European (21) guidelines on the management of endometriosis-associated pain. However, we would encourage investigators to obtain histological confirmation wherever possible.

Thereafter, there was considerable debate about how recently the surgery should have been performed. This proved problematic for three reasons: first, the natural history of endometriosis is uncertain. Second, it might be considered unethical within the context of a trial for medical therapy not to ablate or remove lesions found at an initial laparoscopy given the current recommendation (12). Third, while surgical diagnosis of endometriosis within the last 12

months was considered optimal by some for recruiting into clinical trials, this may actually indicate women having symptoms that persist after surgery (i.e., failed surgical treatment). Thus, a significant proportion of those responding to the meeting summary document considered a surgical diagnosis within the last 5 years to be a more realistic option. A greater proportion of patients in such a group would have a return of symptoms. As many may consider this interval to be too long, it is recommended that a surgical diagnosis should have been made within the last 5 years; however, a shorter interval may be defined. The extent of previous surgical treatment should be documented as this may influence the success of future treatment strategies. This approach assumes that all women with symptoms have endometriosis, even though some may not have evidence of the disease (12).

### 2. What are the appropriate entry criteria?

The only essential entry criteria were a history of endometriosis-associated pain and surgical confirmation of the disease in the chosen time period. These allow a wide variety of treatments to be tested on different stages of disease.

### 3. What are the baseline pain measurements?

Dysmenorrhea and pelvic pain should be measured separately using an 11-point numerical rating scale (NRS), as in other chronic pain conditions (22). These scales should be anchored by 0 = "no pain" and 10 = "worst pain you can imagine." Although in laboratory studies, pain intensity and unpleasantness are measured separately, in clinical trials, using a single measure for each type of pain may be sufficient. Daily pain and the amount of vaginal bleeding should be recorded for at least 1 calendar month before treatment to obtain adequate baseline measurements.

### 4. Should we

- a. Persist with the B&B scale?
- b. Adapt existing pain scale(s)?
- c. Develop a new (patient-derived) pain scale?

The B&B scale was rejected as a primary endpoint but recommended to be retained (albeit in a standardized format) as a secondary endpoint in phase II/III trials. While there was debate as to whether an existing pain scale(s) should be adapted or a new (patient-derived) scale developed, many validated and reproducible measures are already available, and developing a new scale would take a significant amount of time and resources. One group, who are designing a new tool, argued for its utility, although it is not yet in the public domain. Until the new tool can be assessed, validated, and shown to be sensitive to change, the IMMPACT Guidelines (18, 22) should be adapted because of their applicability to other chronic pain conditions.

### 5. Should we take clinical signs into account or not?

Clinical signs should not be used as outcome measures.

### 6. What about comorbidity?

**TABLE 2****Examples of comorbidities to be considered as tertiary endpoints.**

Dyspareunia  
Bladder pain  
Dyschezia  
Fatigue  
Abnormal vaginal bleeding

*Vincent. Pain outcome in endometriosis studies. Fertil Steril 2008.*

Comorbidity was defined as other symptoms associated with endometriosis, rather than separate disease entities such as diabetes. To understand the impact of these symptoms, a woman could choose two or three that bother her the most and regularly report them as tertiary endpoints (Table 2.) For example, for surgical trials involving resection of rectovaginal disease, dyschezia and bowel function are particularly relevant and could be tertiary endpoints.

7. Should we use single (summed) or separate pain assessments?

Separate pain assessments for dysmenorrhea and pelvic pain are superior to a single pain measure, principally because many treatments induce amenorrhea, which makes a summed score unrepresentative of the true extent of improvement. There was disagreement, however, regarding how best to define dysmenorrhea. For example, it was not possible to agree whether pain associated with vaginal spotting at irregular times in the menstrual cycle constitutes dysmenorrhea or not. However, patient representatives felt that most sufferers can tell when their bleeding is menstrual and label this as dysmenorrhea as opposed to pelvic pain associated with spotting. Dysmenorrhea was defined as “pain associated with menstrual bleeding” with the woman’s subjective impression of what menstrual means. The use of a daily diary, to record both pain scores and bleeding, should provide additional support for this approach.

8. Should we measure QoL as well? What about adverse events?

The value of using QoL tools in cancer trials had been questioned; to justify their use as outcome measures, they should detect a change (in either direction) after treatment. Measuring QoL is important and best achieved by using a patient-oriented, disease-specific measure with multiple domains, such as the EHP-30 (16). The EHP-30 measures physical and emotional functioning, two components of the IMMPACT recommendations. Those in favor of QoL measures indicated that surgery for endometriosis can have a profound effect on body image; hormonal treatments can affect emotional functioning. The disease can also impact sexual function, and women have feelings arising from disease-related infertility.

As the EHP-30 appears to assess these aspects of QoL, it should be included. The scores from each domain should

be reported separately in clinical trials rather than as a single overall score, that is, the sum of all the domains.

A tool, such as the Patient Global Impression of Change (23), could be used for adverse event reporting. Although it is still a part of IMMPACT, it is no longer in common use, and it is not part of these recommendations.

9. How often should we measure?

Pain and bleeding should be reported on daily at a consistent time. Patients may prefer to do this in the evenings, as mornings tend to be too busy. Some suggested the use of electronic diaries or Web-based scoring systems, but these may not be applicable to international studies.

The frequency and time intervals for the measurement of secondary endpoints such as EHP-30 scores were discussed. Transient side effects and/or early symptomatic improvements might be missed if data collection was not well timed. These should be measured weekly for the first 6 weeks, then monthly until 6 months, and then at 9, 12, 18, and 24 months. Longer term data (e.g., up to 5 years post-treatment) would be valuable and should be collected when possible.

As many possible tertiary endpoints may have a cyclical component, these should be scored daily similar to dysmenorrhea and pelvic pain. A “not applicable” box for symptoms such as dyspareunia and dyschezia that are not present should be an option as a score of 0 could be misleading and suggest improvement.

10. How do we address cyclicity?

The daily collection of information on both pain and bleeding was felt to be sufficient to capture cyclicity.

11. Which rescue medications should be allowed/recorded and how?

The use of rescue analgesia and complementary therapies should be captured. The pain score immediately before the use of such treatments should be recorded, as well as the indication, that is, were they being taken for endometriosis-associated pain or an unrelated symptom? This would also allow information on adverse events to be captured if, for example, the analgesia was required for a headache that was secondary to hormonal treatment. A longitudinal analysis of rescue medication should also give an idea of the timescale required for the treatment to have an effect. Restricting the use of rescue medication would be both unethical and likely to increase the number of dropouts from the trials.

12. How do we define a responder?

A clear definition of a responder should be provided in each trial. It is suggested that this be either a >30% or >50% reduction in symptoms; however, the precise definition will depend on the trial.

13. What is a clinically meaningful effect?

The definition of a clinically meaningful effect should be patient determined.

## SPECIFIC RECOMMENDATIONS

The specific recommendations of the Art and Science of Endometriosis meeting were as follows:

1. Entry criteria
  - a. Surgical diagnosis of endometriosis within the last 5 years
  - b. Pain symptoms
  - c. Data capture at baseline:
    - i. ASRM classification (24)
    - ii. Baseline pain scores over at least two menstrual cycles
    - iii. EHP-30 (16)
  - d. Previous treatments and responses
2. Primary outcome measures
  - a. Daily ratings of pelvic pain
  - b. Daily ratings of dysmenorrhea
  - c. Ratings on an 11-point NRS, anchored by 0 = “no pain” and 10 = “worst pain you can imagine” based on a recall of the worst pain experienced over the previous 24 hours. Daily record of bleeding as none, spotting, light, or heavy compared with a normal period.
3. Secondary outcome measures
  - a. B&B (2) with separate scores for each domain, administered weekly for 6 weeks, then monthly until 6 months, then at 9, 12, 18, and 24 months
  - b. EHP-30 (16) with separate and total scores, administered at the same time points as the B&B
  - c. Use of rescue analgesia/therapies including an NRS before use and a record of the indication
  - d. Study-specific adverse event questionnaires with direct questions and free text, administered at the same time points as the B&B
  - e. Detailed information as per the CONSORT guidelines (25), including
    - i. The recruitment process
    - ii. The number of candidate participants who were excluded and why
    - iii. The number of candidates who chose not to enter the trial and why
    - iv. The use of prohibited concomitant medications and other protocol deviations
    - v. The number and reasons for withdrawal from each treatment group
    - vi. The types, rates, and reasons for nonadherence with treatment in each group
4. Tertiary outcome measures

Daily NRS (or Not Applicable [NA]) of three symptoms the patient feels are important to her, for example, dyspareunia, dyschezia, fatigue, and so on.

## CONCLUSION

The Art and Science of Endometriosis meeting provided a forum for clinicians and scientists from around the world to discuss appropriate entry criteria and outcome measures for

clinical trials in pain related to endometriosis. Adapting current recommendations from other chronic pain conditions reduces the unnecessary waste of time and resources associated with the development of new tools. However, the use of disease-specific and patient-centered measures should allow clinically relevant and useful information to be gathered. It is hoped that these suggestions are adopted in future trials such that evidence-based practice can increase in this difficult area.

*Acknowledgments:* The authors thank the following people for presenting at the meeting and facilitating later discussions: Karen Berkley, Ph.D., Florida State University, Tallahassee; David Olive, M.D., University of Wisconsin-Madison; Paolo Vercillini, M.D., University of Milan; Crispin Jenkinson, Ph.D., University of Oxford; MaryLou Ballweg, Endometriosis Association; Charles Cleeland, Ph.D., UT MD Anderson Cancer Center; Charlie Cao, Ph.D., Takeda Global Research and Development Center.

## REFERENCES

1. Streiner DL, Norman G. Health Measurement Scales—a practical guide to their development and use. 3d ed. Oxford: Oxford University Press, 2003.
2. Biberoglu KO, Behrman SJ. Dosage aspects of danazol therapy in endometriosis: short term and long term effectiveness. *Am J Obstet Gynecol* 1981;139:645.
3. Lemay A, Maheux R, Hout C, Blanchet J, Faure N. Efficacy of intranasal or subcutaneous luteinizing hormone–releasing hormone agonist inhibition of ovarian function in the treatment of endometriosis. *Am J Obstet Gynecol* 1988;158:233–6.
4. Hornstein MD, Yuzpe AA, Burry KA, Heinrichs LR, Buttram-VL J, Orwoll ES. Prospective randomized double-blind trial of 3 versus 6 months of nafarelin therapy for endometriosis associated pelvic pain. *Fertil Steril* 1995;63:955–62.
5. Shaw RW. An open randomized comparative study of the effect of goserelin depot and danazol in the treatment of endometriosis. *Fertil Steril* 1992;58:265–72.
6. Vercellini P, Fedele L, Aimi G, Pietropaolo G, Consonni D, Crosignani PG. Association between endometriosis stage, lesion type, patient characteristics and severity of pelvic pain symptoms: a multivariate analysis of over 1000 patients. *Hum Reprod* 2007;22:266–71.
7. Szendei G, Hernadi Z, Devenyi N, Csapo Z. Is there any correlation between stages of endometriosis and severity of chronic pelvic pain? Possibilities of treatment. *Gynecol Endocrinol* 2005;21:93–100.
8. Mahmood T, Templeton A, Thomson L, Fraser C. Menstrual symptoms in women with pelvic endometriosis. *Br J Obstet Gynecol* 1991;98:558–63.
9. Anaf V, Simon P, El Nakadi I, Fayt I, Buxant F, Simonart T, et al. Relationship between endometriotic foci and nerves in rectovaginal endometriotic nodules. *Hum Reprod* 2000;15:1744–50.
10. Berkley KJ, Rapkin AJ, Papka RE. The pains of endometriosis. *Science* 2005;308:1587–9.
11. Bodner CH, Garratt AM, Ratcliffe J, Macdonald LM, Penney G. Measuring health-related quality of life outcomes in women with endometriosis—results of the Gynaecology Audit Project in Scotland. *Health Bull (Edinb)* 1997;55:109–17.
12. Abbott JA, Hawe J, Clayton RD, Garry R. The effects and effectiveness of laparoscopic excision of endometriosis: a prospective study with 2–5 year follow-up. *Hum Reprod* 2003;18:1922–7.
13. Marques TEA, Bahamondes L, Aldrighi JM, Petta CA. Quality of life in Brazilian women with endometriosis assessed through a medical outcome questionnaire. *J Reprod Med* 2004;49:115–20.
14. Jones G, Jenkinson C, Kennedy S. The impact of endometriosis upon quality of life: a qualitative analysis. *J Psycho Som Obs Gyn* 2004;25:123–33.
15. Jones G, Jenkinson C, Kennedy S. Development of the Short Form Endometriosis Health Profile Questionnaire: the EHP-5. *Qual Life Res* 2004;13:695–704.

16. Jones G, Kennedy S, Barnard A, Wong J, Jenkinson C. Development of an Endometriosis Quality-of-Life Instrument: The Endometriosis Health Profile-30. *Obstet Gynecol* 2001;98:258–64.
17. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med Sing* 1994;23:129–38.
18. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2003;106:337–45.
19. Stegmann BJ, Sinaii N, Liu S, Segars J, Merino M, Nieman LK, et al. Using location, color, size, and depth to characterize and identify endometriosis lesions in a cohort of 133 women. *Fertil Steril* 2008;89:1632–6.
20. American Society for Reproductive Medicine. Treatment of pelvic pain associated with endometriosis. *Fertil Steril* 2006;86:S18–27.
21. Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, et al. ESHRE guideline for the diagnosis and treatment of endometriosis. *Hum Reprod* 2005;20:2698–704.
22. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;113:9–19.
23. Guy W. ECDEU assessment manual for psychopharmacology. Washington, D.C.: U.S. Government Printing Office, 1976.
24. American Society for Reproductive Medicine. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertil Steril* 1997;67:817–21.
25. Moher D, Schulz KF, Altman DG, for the CG. The CONSORT Statement: Revised Recommendations for Improving the Quality of Reports of Parallel-Group Randomized Trials. *Ann Intern Med* 2001;134:657–62.