

Published in final edited form as:

*J Clin Epidemiol.* 2011 March ; 64(3): 293–300. doi:10.1016/j.jclinepi.2010.03.017.

## Systematic review highlights difficulty interpreting diverse clinical outcomes in abnormal uterine bleeding trials

David D. Rahn<sup>a,\*</sup>, Husam Abed<sup>b</sup>, Vivian W. Sung<sup>c</sup>, Kristen A. Matteson<sup>c</sup>, Rebecca G. Rogers<sup>d</sup>, Michelle Y. Morrill<sup>e</sup>, Matthew D. Barber<sup>f</sup>, Joseph I. Schaffer<sup>a</sup>, Thomas L. Wheeler II<sup>g</sup>, Ethan M. Balk<sup>h</sup>, and Katrin Uhlig<sup>h</sup> for the Society of Gynecologic Surgeons—Systematic Review Group

<sup>a</sup>University of Texas Southwestern Medical Center, Dallas, TX, USA

<sup>b</sup>Henry Ford Health System, Detroit, MI, USA

<sup>c</sup>Women and Infants Hospital of Rhode Island/Brown Medical School, Providence, RI, USA

<sup>d</sup>University of New Mexico Health Sciences Center, Albuquerque, NM, USA

<sup>e</sup>Kaiser Permanente, San Francisco, CA, USA

<sup>f</sup>Cleveland Clinic, Cleveland, OH, USA

<sup>g</sup>University of South Carolina, Greenville, SC, USA

<sup>h</sup>Tufts Medical Center, Boston, MA, USA

### Abstract

**Objectives**—(1) To systematically collect and organize into clinical categories all outcomes reported in trials for abnormal uterine bleeding (AUB); (2) to rank the importance of outcomes for patient decision making; and (3) to improve future comparisons of effects in trials of AUB interventions.

**Study Design and Setting**—Systematic review of English-language randomized controlled trials of AUB treatments in MEDLINE from 1950 to June 2008. All outcomes and definitions were extracted and organized into major outcome categories by an expert group. Each outcome was ranked “critically important,” “important,” or “not important” for informing patients’ choices.

**Results**—One hundred thirteen articles from 79 trials met the criteria. One hundred fourteen different outcomes were identified, only 15 (13%) of which were ranked as critically important and 29 (25%) as important. Outcomes were grouped into eight categories: (1) bleeding; (2) quality of life; (3) pain; (4) sexual health; (5) patient satisfaction; (6) bulk-related complaints; (7) need for subsequent surgical treatment; and (8) adverse events.

**Conclusion**—To improve the quality, consistency, and utility of future AUB trials, we recommend assessing a limited number of clinical outcomes for bleeding, disease-specific quality of life, pain, sexual health, and bulk-related symptoms both before and after treatment and reporting satisfaction and adverse events. Further development of validated patient-based outcome measures and the standardization of outcome reporting are needed.

© 2011 Elsevier Inc. All rights reserved.

\*Corresponding author. 5323 Harry Hines Boulevard, Dallas, TX 75390-9032, USA. Tel.: +214-648-6430; fax: +214-648-0328. David.Rahn@utsouthwestern.edu (D.D. Rahn).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jclinepi.2010.03.017.

## Keywords

Menorrhagia; Dysfunctional uterine bleeding; Leiomyomata; Hysterectomy; Uterine artery embolization; Endometrial ablation

---

## 1. Introduction

Abnormal uterine bleeding (AUB), an alteration in the volume, pattern, or duration of menstrual blood flow [1], is the single, most common reason for gynecologic referral [2]. The large societal and personal burden of AUB lies in its major impact on quality of life, productivity, health care use, and costs [3,4]. Although hysterectomy cures AUB, it carries the risks associated with major surgery. A multitude of alternative less-invasive procedures have gained in popularity, including endometrial destruction techniques, myomectomy, polypectomy, and uterine artery embolization. Medical management of AUB includes modification of endometrial coagulation factors or prostaglandins, hormonal manipulation, and insertion of levonorgestrel-releasing intrauterine systems. This myriad of therapeutic options makes direct comparisons of the effectiveness of different AUB interventions necessary but difficult.

Further complicating meaningful comparisons of interventions for AUB is the subjective and variable nature of the presenting symptoms. Nearly half the women who seek treatment for AUB lose less blood monthly than the 80 mL that strictly defines “heavy menstrual bleeding” [5,6]. Rather, many patients may be more bothered by the unpredictability of bleeding or pain associated with passage of large clots. As such, outcomes that capture the symptoms most relevant to patients (so-called patient-reported outcome measures) are increasingly being used in research on treatments for AUB [1]. In this study, our objectives were (1) to create a complete inventory of outcomes reported in published randomized controlled trials (RCTs) for treatment of AUB and organize them into overarching categories, and (2) to rank the importance of these measures by a physician expert-review process that considers the clinical relevance of the outcome to the patient and the quality of the measuring instrument. The ultimate goal of this review was to facilitate and improve future comparisons between different interventions for AUB.

## 2. Methods

### 2.1. Study sponsor and investigators

The Systematic Review Group (SRG) of the Society of Gynecologic Surgeons (SGS) is a group of 25 physicians with expertise in benign gynecology (including AUB), urogynecology, obstetrics, and consultants with expertise in systematic review methods and guideline development. Members and relevant disclosures are described elsewhere (Appendix A, available on the journal’s Web site at [www.elsevier.com](http://www.elsevier.com)).

### 2.2. Study eligibility and search strategy

The SRG undertook a systematic search to identify RCTs of treatments for AUB in women with either dysfunctional uterine bleeding (DUB) or fibroids. We performed a MEDLINE search of articles published between 1950 and June 11, 2008. Search terms included menorrhagia, menometrorrhagia, uterine hemorrhage, dysfunctional uterine bleeding, abnormal uterine bleeding, heavy vaginal bleeding, fibroid, leiomyoma, anovulatory bleeding, hysterectomy, myomectomy, uterine artery embolization, endometrial ablation, intrauterine device or system, several classes of medical therapy, and a search module for RCTs. The search was limited to human and English-language studies. A manual review of

bibliographies of retrieved articles was completed to identify additional studies. Inclusion criteria required RCTs to have at least 10 individuals per arm and at least 1-month duration of follow-up. We excluded RCTs of “pretreatments” before a subsequent intervention and analyses assessing only cost or resource use.

### 2.3. Data extraction

Studies meeting criteria were independently data extracted in duplicate by members of the review group, most of whom had experience with data extraction from a prior systematic review project [7]. The data fields included study size; duration of follow-up; population (DUB, fibroids, or mixed); interventions compared; and outcomes (including patient-reported outcome measures and adverse events). A complete inventory of all reported outcomes was created. Each individual outcome was detailed in terms of its specific definition, the time point(s) when it was assessed, which instrument or test was used to measure it, whether the instrument or test had been validated, the units or metric used for presenting the results, and whether it was the “primary outcome” of the study (either explicitly described as such or used in a power calculation to determine study size). Composite outcomes were designated as such. All outcomes from the same trial were recorded on a single data extraction form. The list of all unique outcomes comprised the “outcome inventory.”

### 2.4. Overarching outcome categories

From this outcome inventory, the outcomes were organized and grouped into eight proposed overarching outcome domains. Categories were determined based on their applicability to all potential interventions for AUB and the physician expert group’s consensus of their relevance for informing patient choices. Outcomes related to cost, resource use, or those determined by the review group to have limited relevance for assessing clinical effectiveness were excluded from categorization and further analyses.

### 2.5. Assessing importance of outcomes

In preparation for ranking the importance of outcomes, the group completed a review of AUB outcome measures. The review included reports on the subjective experience of patients with AUB [8-13] and evaluations of the quality of patient-based outcome measures used for AUB, including findings from focus groups conducted by one of the coauthors (K.A.M.) on what women with AUB find important or bothersome [1,14]. The importance of each outcome was graded on a scale suggested by the GRADE working group (Grading of Recommendations Assessment, Development, and Evaluation) [15]. This working group, which strives to standardize and improve the transparency of grading evidence and recommendations in guideline development, proposes a 9-point scale with outcomes scored as “critical” for decision making (score: 7–9), “important but not critical” (score: 4–6), and “not important” for decision making (score: 1–3). The GRADE group does not provide specific details for what to consider in the ranking. Our group had several discussions on considerations potentially impacting the importance of an outcome at in-person meetings or on conference calls. There was general agreement to consider the magnitude of the impact of the outcome on a patient’s well-being and the quality of the measurement tool used to assess the outcome in terms of what was known about its validity, reliability, and feasibility. The overall impact of the outcome on patient’s health and well-being superseded the assessment of the instrument used to measure the outcome. Consideration was given to whether the instruments were multi- or unidimensional, generic or condition specific, or psychometrically tested. The rankings for outcomes were increased if they were measured using multidimensional, condition-specific, or psychometrically tested instruments. Psychometrically robust instruments were thought by the group to have the capacity to more broadly capture symptoms important from the patients’ perspective (multidimensional or

condition specific) or the ability to provide higher-quality measurement of outcomes (psychometrically tested). The group followed an iterative process of consensus building.

A ballot system was used by all members to grade the importance of each outcome based on his or her qualitative synthesis of the aforementioned considerations. The mean was calculated and rounded to the nearest whole number. Members were asked to report any potential conflicts of interest for this vote related to professional practice, research activities, or financial investments, and, if present, to abstain from voting. No conflicts were reported. For outcomes grouped under adverse events, no score was tallied as the group adapted the Consolidated Standards of Reporting Trials (CONSORT) statement on reporting of harms [16].

Finally, the group developed recommendations for outcome measures to assess in future trials for AUB based on the outcome inventory, the overarching outcome categories, and the “importance” rating of individual outcomes.

### 3. Results

#### 3.1. Search yield and outcome inventory

The literature search identified 5,025 citations. Of these, 4,892 were excluded after review of title or abstract for not being treatment trials for AUB. After examining the full text of 133 articles, 20 were excluded for not being RCTs or if they were studies of resource use and cost estimations (Fig. 1). Studies ranged in size from 20 to 372 patients. Length of follow-up ranged from only perioperative events to 5 years of follow-up. Appendix B (available on the journal’s Web site at [www.elsevier.com](http://www.elsevier.com)) lists detailed summary data for the final 79 trials reported in 113 articles, of which 51 trials involved patients with DUB (72 articles), 20 trials with uterine fibroids contributing to abnormal bleeding (27 articles), and eight trials with mixed or unknown/unclassified cause of AUB (14 articles). Of the DUB trials, seven were surgical trials that included hysterectomy as an intervention. Less-invasive surgical interventions involving endometrial ablation or resection were reported in 31 trials. The remainder of DUB trials (19 articles) examined the role of medications, largely nonsteroidal anti-inflammatory and hormonal therapies. Most trials reporting on treatments for fibroid-associated AUB were surgical, with six involving hysterectomy. Myomectomy and uterine artery embolization were used in four and eight fibroid trials, respectively.

The 79 trials contained 114 outcomes, including 27 unique adverse events. After review of the inventory of outcomes, the SRG experts deemed 43 outcomes (38%) as not important for long-term clinical decision making (Appendix C [available on the journal’s Web site at [www.elsevier.com](http://www.elsevier.com)]), for example, estradiol and ferritin levels, grading of uterine artery spasm, fibroid volume, and volume of embolic material. There was no further voting on adverse event outcomes, as it was agreed that these should be reported consistent with the recommendations of the CONSORT statement [16]; examples of these events include death, pulmonary embolism, wound hematoma, and fluid overload. Of the remaining 44 outcomes, the most commonly reported outcomes were patient satisfaction (35 of 79 trials, 44%), rate of amenorrhea (32 of 79 trials, 41%), and frequency of subsequently needed interventions—either hysterectomy (25 of 79 trials, 32%) or other nonhysterectomy surgical treatment (26 of 79 trials, 33%).

#### 3.2. Outcome categories and ranking of importance

The overarching categories agreed upon by the SRG were (1) bleeding; (2) quality of life; (3) pain; (4) sexual health; (5) bulk-related symptoms (for patients with anatomic/fibroid-related AUB); (6) patient satisfaction; (7) need for additional treatments; and (8) adverse events. All outcomes deemed either “important” or “critical” for clinical decision making for

treating AUB were grouped into these eight categories. Table 1 shows a listing of these outcomes with their corresponding scores.

When rating an outcome, we decided that it was appropriate to consider *the concepts* that an individual instrument measured. The group agreed on some principles. A comprehensive and multifaceted instrument rated higher than a focused unidimensional instrument, because the former has the capacity to more broadly capture symptoms that would be important from the patients' perspectives. Condition-specific quality-of-life instruments rated higher than general instruments, because they had the capacity to include more meaningful questions related to the intermittent symptoms of AUB. In the bleeding category, for example, instruments that can assess both heaviness and irregularity of bleeding were given the highest scores and were most likely to be considered "critical" by our group. The Pictorial Bleeding Assessment Chart is one such measure, and it was commonly used ( $n = 22$  trials). Other tools that examined just heaviness of bleeding were considered important but not critical, for example, sanitary pad counting ( $n = 7$  trials).

#### 4. Discussion

Given the prevalence of AUB, high-quality research on treatment modalities for this burdensome condition is imperative for the advancement of women's health-related quality of life [17-19]. In our systematic review, we identified 79 trials for AUB that reported 114 different outcomes, many of which attempted to evaluate outcomes from the patient perspective, a top priority in health research. However, the sheer number and variety of outcomes reported in the literature hinders comparing different interventions across different studies. This limits the utility of this literature for informing clinicians, patients, and policy makers about the respective benefits and harms of different interventions. In response to this dilemma, we generated an outcomes inventory and ranked the importance of outcomes in an explicit expert decision-making process to guide evidence synthesis and future research.

The outcome inventory generated included all clinically relevant patient-reported outcomes that have been used in *past* studies for AUB over the past 58 years. Categorizing the individual outcomes into overarching outcome domains and then rating each individual outcome in terms of "importance" for clinical decision making provides a point of reference for recommendations of outcomes to be used in *future* studies on AUB. Our ranking emphasizes patient-reported outcomes deemed to be important but not explicitly how to measure them, as currently, there is no definitive answer on how best to measure all of these outcomes. The variety of outcomes assessed and instruments used in our review highlights that there is no high-quality, standardized set of instruments that comprehensively evaluates outcomes for women with AUB that is universally accepted.

Although reliable, validated instruments exist for the specific determination of the amount of blood loss [20,21]; there is no single instrument that is considered the standard for evaluating menstrual symptoms, AUB-related quality of life, sexual health, or satisfaction with treatment in women with AUB. Further research is needed in the development of AUB-related quality-of-life measures that are proven responsive to capture changes after treatment. Generic health-related quality-of-life questionnaires can be used to augment validated disease-specific measures of blood loss; however, some women with AUB may experience difficulty in answering generic questions about health perceptions because of the intermittent nature of AUB [14,22-24]. As AUB symptoms can be chronic and intermittent, it is important to assess more than a single aspect of a woman's bleeding experience, and that is why we have included all outcomes categories rated as "important" by our review group in the following recommendations.

Specification of important outcomes is a key step when developing a systematic review protocol to ensure that useful information is synthesized. The recently updated Cochrane Handbook for Systematic Reviews of Interventions recommends choosing a maximum of seven outcomes and to present the main findings on those in the “Summary of findings” table [25]. This is because evidence synthesis across an excessive number of outcomes becomes confusing. The eight outcome categories proposed for AUB trials are not necessarily independent but capture different aspects of a disease that manifests with varying symptoms that can have variable impact on a woman’s health. Hopefully, going forward, single instruments will collapse the measurement of symptoms and manifestations across different categories, but at this point, the eight outcome categories are proposed as a first step to help with cleaning up a field plagued by a plethora of outcomes.

The problems encountered in this systematic review of AUB trials are not unique to gynecology. Similar strides for developing meaningful and parsimonious outcome measures have been undertaken in other medical disciplines. For example, OMERACT (Outcome Measures in Rheumatoid Arthritis Clinical Trials) has endeavored to improve national and international consensus on issues, such as minimum number of outcome measures to be included in treatment trials, and to decide on the magnitude of differences judged to be clinically relevant [26].

#### 4.1. Recommendations for outcomes in abnormal uterine bleeding trials

The following recommendations provide guidance for a more standardized assessment and reporting of outcomes specifically for trials for AUB. Our recommendations should be considered supplemental to the CONSORT guidelines [27] and the extensions of the CONSORT guidelines for trials assessing nonpharmacological treatments [28] and for assessing the harms [16].

To fully assess the effects of an intervention, we recommend assessing outcomes in all eight outcome categories (Table 2). Assessments of bleeding, quality of life, pain related to heavy bleeding, sexual health, and bulk-related complaints (in patients with leiomyomas) should be conducted both before and after treatment, and the change for each patient should be used as an analytic metric. Patient satisfaction, need for additional treatment, and adverse events should also be reported.

The goal of any treatment for AUB is to reduce or eliminate bleeding, improve other associated symptoms, and improve quality of life. We suggest that bleeding be assessed based on patient-reported amount, frequency of menses, duration of bleeding, and regularity of menses [29]. Many studies include pain, an important symptom in terms of potential impact on patients’ health and well-being, as an outcome, and this should also be measured. Women with leiomyomata may also have symptoms specific to uterine size. Therefore, trials in women with leiomyomata should assess bulk-related symptoms as an outcome before and after treatment.

Condition-specific quality of life is a key patient-based outcome measure. The large personal burden of AUB lies in its adverse impact on quality of life [3,4], and this is an important outcome for AUB research. Sexual health, another important part of a woman’s quality of life, may be adversely affected by AUB and may be affected differentially by different treatments for AUB. Thus, sexual health should also be assessed before and after treatment.

Many studies we reviewed included “satisfaction with treatment” as an outcome for AUB treatments. We rated satisfaction with treatment as an important outcome, given its face validity, despite the dependency of this outcome on a patient’s pretreatment expectations

and goals. Measured as the sole outcome for a study, “satisfaction with treatment” is limited, as it provides no information in terms of which symptoms the treatment alleviated and why the patient is satisfied. Collected as one of several outcomes, satisfaction with treatment may still be helpful to inform patients in their selection of different therapies.

These recommendations are not meant to provide an exhaustive or definitive list of the exact data and outcome measures to be collected. There is no one standard measurement tool for the suggested outcomes included in these recommendations. In the absence of a standardized instrument for measuring outcomes in trials for AUB, researchers should consider combining available validated instruments for measuring outcomes or clearly explain how new instruments for measuring outcomes were developed specifically for their study. Researchers should clearly describe which instruments or methods were used to assess bleeding, related symptoms, quality of life, sexual health, and satisfaction. It would be ideal if investigators developed a validated measure or set of instruments that captured all AUB-related symptoms so that a standardized tool can be used in all future trials.

#### 4.2. Strengths and limitations

One strength of this study is that the review group used a systematic process to identify RCTs for AUB and to list all outcomes reported in these trials. Additionally, a data-driven iterative expert consensus process was used to grade the importance of each of these outcomes. We used a voting process that left it to each group member to weigh the relative impact of the various considerations for what makes an outcome important for decision making and to arrive at a composite grade. Because this was, to some degree, subjective, we would have excluded individuals with any self-reported conflicts of interest, had there been any. As the GRADE working group states, knowledge about optimal strategies for making decisions about relative importance of outcomes remains limited. Regardless, it is desirable to make the process transparent and explicit [15]. A limitation of the study is that there was no direct patient participation in deciding on the importance of outcomes for clinical decision making. For this reason, we referred to the available literature on patient experience with AUB and patient-based outcomes measures for AUB. For the same reason, we invited individuals with particular expertise in patient-reported outcome measures for AUB research to participate in the review group discussions.

### 5. Conclusion

An explicit, transparent process for grading the importance of outcomes is a key step in evidence synthesis and assessment of benefits and limitations of alternative treatments. Our process of ranking importance of AUB outcomes combined systematic review steps to generate an outcome inventory followed by a stepwise expert-based categorization and rating of outcomes. We have provided recommendations for future AUB trials to improve the consistency and quality of outcome data. This process for ranking of importance of outcomes may be used in other fields where heterogeneity in outcomes hampers comparison of results across different treatments.

To fundamentally address and ameliorate the problem of nonstandardized outcome measurement for AUB will require further work involving clinicians, patients, and researchers to narrow the range of outcomes to a core set of valid, discriminatory, and feasible tools and to use them consistently in future trials.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The following SGS Systematic Review Group members significantly contributed to the acquisition of data: LTC Jeffrey L. Clemons, MD; Rajiv B. Gala, MD; Oz Harmanli, MD; Lior Lowenstein, MD, MS; James Lukban, DO; Abraham Morse, MD; Miles Murphy, MD, MSPH; Nicole Calloway Rankins, MD, MPH; Scott Smilen, MD; Karl Tamussino, MD; and James Theofrastous, MD.

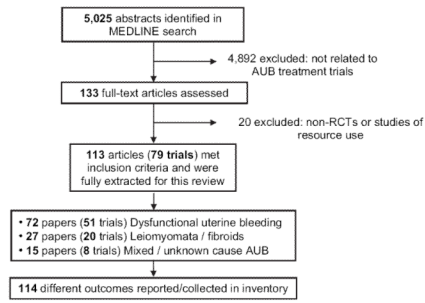
Funding/support: Funding of assistance by methodologic experts in systematic review and logistic support was provided by the Society of Gynecologic Surgeons.

## References

1. Matteson KA, Boardman LA, Munro MG, Clark MA. Abnormal uterine bleeding: a review of patient-based outcome measures. *Fertil Steril* 2009;92:205–16. [PubMed: 18635169]
2. van Dongen H, van de Merwe AG, de Kroon CD, Jansen FW. Diagnostic hysteroscopy in abnormal uterine bleeding: a systematic review and meta-analysis. *J Minim Invasive Gynecol* 2009;16:47–51. [PubMed: 18990612]
3. Côté I, Jacobs P, Cumming DC. Use of health services associated with increased menstrual loss in the United States. *Am J Obstet Gynecol* 2003;188:343–8. [PubMed: 12592237]
4. Liu Z, Doan QV, Blumenthal P, Dubois RW. A systematic review evaluating health-related quality of life, work impairment, and health-care costs and utilization in abnormal uterine bleeding. *Value Health* 2007;10:183–94. [PubMed: 17532811]
5. Haynes P, Hodgson H, Anderson AB, Turnbull AC. Measurement of menstrual blood loss in patients complaining of menorrhagia. *Br J Obstet Gynaecol* 1977;84:763–8. [PubMed: 921913]
6. Warner PE, Critchley HO, Lumsden MA, Campbell-Brown M, Douglas A, Murray GD. Menorrhagia I: measured blood loss, clinical features, and outcome in women with heavy periods: a survey with follow-up data. *Am J Obstet Gynecol* 2004;190:1216–23. [PubMed: 15167821]
7. Sung VW, Rogers RG, Schaffer JI, Balk EM, Uhlig K, Lau J, et al. Graft use in transvaginal pelvic organ prolapse repair: a systematic review. *Obstet Gynecol* 2008;112:1131–42. [PubMed: 18978116]
8. O'Flynn N, Britten N. Menorrhagia in general practice—disease or illness. *Soc Sci Med* 2000;50:651–61. [PubMed: 10658846]
9. O'Flynn N. Menstrual symptoms: the importance of social factors in women's experiences. *Br J Gen Pract* 2006;56:950–7. [PubMed: 17132384]
10. Santer M, Wyke S, Warner P. What aspects of periods are most bothersome for women reporting heavy menstrual bleeding? Community survey and qualitative study. *BMC Womens Health* 2007;7:8. [PubMed: 17543131]
11. Santer M, Wyke S, Warner P. Women's management of menstrual symptoms: findings from a postal survey and qualitative interviews. *Soc Sci Med* 2008;66:276–88. [PubMed: 17920178]
12. Warner P, Critchley HOD, Lumsden MA, Campbell-Brown M, Douglas A, Murray G. Referral for menstrual problems: cross-sectional survey of symptoms, reasons for referral, and management. *BMJ* 2001;323:24–8. [PubMed: 11440940]
13. Garside R, Britten N, Stein K. The experience of heavy menstrual bleeding: a systematic review and meta-ethnography of qualitative studies. *J Adv Nurs* 2008;63:550–62. [PubMed: 18808575]
14. Clark TJ, Khan KS, Foon R, Pattison H, Bryan S, Gupta JK. Quality of life instruments in studies of menorrhagia: a systematic review. *Eur J Obstet Gynecol Reprod Biol* 2002;104:96–104. [PubMed: 12206918]
15. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. GRADE Working Group. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008;336:995–8. [PubMed: 18456631]
16. Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781–8. [PubMed: 15545678]
17. Barnard K, Frayne SM, Skinner KM, Sullivan LM. Health status among women with menstrual symptoms. *J Womens Health* 2003;12:911–9.



18. Kjerulff KH, Erickson BA, Langenberg PW. Chronic gynecological conditions reported by US women: findings from the National Health Interview Survey, 1984 to 1992. *Am J Public Health* 1996;86:195–9. [PubMed: 8633735]
19. Cote I, Jacobs P, Cumming DC. Use of health services associated with increased menstrual loss in the United States. *Am J Obstet Gynecol* 2003;188:343–8. [PubMed: 12592237]
20. Higham JM, O'Brien PM, Shaw RW. Assessment of menstrual blood loss using a pictorial chart. *Br J Obstet Gynaecol* 1990;97:734–9. [PubMed: 2400752]
21. Hallberg L, Nilsson L. Determination of menstrual blood loss. *Scand J Clin Lab Invest* 1964;16:244–8. [PubMed: 14161862]
22. Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT. The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *BMJ* 1993;306:1440–4. [PubMed: 8518640]
23. Garratt AM, Ruta DA, Abdalla MI, Russell IT. SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. *Qual Health Care* 1994;3:186–92. [PubMed: 10140232]
24. Jenkinson C, Peto V, Coulter A. Making sense of ambiguity: evaluation in internal reliability and face validity of the SF 36 questionnaire in women presenting with menorrhagia. *Qual Health Care* 1996;5:9–12. [PubMed: 10157276]
25. Higgins, JPT.; Green, S., editors. *Cochrane handbook for systematic reviews of interventions version 5.0.2 [updated September 2009]*. The Cochrane Collaboration. 2009 [March 6, 2010]. Available at [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
26. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007;8:38. [PubMed: 18039364]
27. Moher D, Schulz KF, Altman D for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285:1987–91. [PubMed: 11308435]
28. Boutron I, Moher D, Altman D, Schulz K, Ravaud P for the CONSORT Group. Methods and processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments. *Ann Intern Med* 2008;148:W-60–6. [PubMed: 18283201]
29. Fraser IS, Critchley HO, Munro MG, Broder M. Can we achieve international agreement on terminologies and definitions used to describe abnormalities of menstrual bleeding. *Hum Reprod* 2007;22:635–43. [PubMed: 17204526]



**Fig. 1.** Flow diagram of systematic review of randomized controlled trials (RCTs) for abnormal uterine bleeding (AUB).

**Table 1**

Frequencies of outcomes across RCTs for AUB and their importance in clinical decision making

Outcome	Frequency of outcome reporting, <i>N</i> (within 79 RCTs for AUB)	GRADE score (1–9)	Outcome importance
<b>Bleeding</b>			
Pictorial diary (PBAC)	22	8	C
Bleeding diary	4	8	C
Duration of menses/duration of cycle	20	7	C
Bleeding score	6	8	C
Blood loss evaluation with sanitary napkins with use of alkaline haematin at 550nm	15	5	I
Menstrual changes (amenorrhea, hypomenorrhea, eumenorrhea)	25	7	C
Total menstrual fluid loss	5	5	I
Sanitary-towel count	7	6	I
Monthly blood loss index	2	7	C
Amenorrhea	32	7	C
<b>QOL</b>			
Menstrual impact questionnaire	1	7	C
Menorrhagia QOL	2	7	C
Fibroid QOL	3	7	C
SF-36	19	6	I
SF-12	3	5	I
Euro QOL-5d	5	5	I
Psychosocial adjustment to illness scale	2	5	I
Health utility Index Mark-3	1	5	I
Spielberger State-Trait Anxiety Inventory	2	4	I
Hospital Anxiety and Depression Scale	3	4	I
Beck's depression inventory	1	4	I
Eysenck personality questionnaire	1	4	I
<b>Pain</b>			
Pelvic pain	19	7	C
Dysmenorrhea	21	7	C
Abdominal pain	5	6	I
Low-back pain	5	5	I
<b>Sexual health</b>			
Return to sexual activity	5	5	I
Sexual activity	6	5	I
Body image/sexuality	2	6	I
Revised Sabbatsberg Sexual Rating Scale	1	5	I
McCoy Sex Scale	1	5	I
Golombok Rust inventory of marital state	1	5	I
Decreased libido	3	4	I

Outcome	Frequency of outcome reporting, <i>N</i> (within 79 RCTs for AUB)	GRADE score (1–9)	Outcome importance
Dyspareunia	8	5	I
Sexual satisfaction	4	5	I
Bulk-related symptoms			
Pelvic pressure—fibroids	9	8	C
Pelvic pressure—DUB	1	6	I
Defecatory dysfunction—fibroids	2	5	I
Defecatory dysfunction—DUB	1	4	I
Bowel symptoms—fibroids	1	5	I
Bowel symptoms—DUB	2	4	I
Urinary symptoms—fibroids	5	5	I
Urinary symptoms—DUB	3	4	I
Patient satisfaction			
Satisfaction	35	6	I
Recommending the treatment to a friend	8	6	I
Additional treatments			
Hysterectomy	25	8	C
Other than hysterectomy	26	7	C

*Abbreviations:* RCT, randomized controlled trial; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; PBAC, Pictorial Blood Assessment Chart; QOL, quality of life; SF-36 and SF-12, Medical Outcomes Study—Short-Form 36 and Short-Form 12; DUB, dysfunctional uterine bleeding; C, “critical” for clinical decision making—GRADE group nomenclature (scores of 7–9); I, “important” for clinical decision making—GRADE group nomenclature (scores of 4–6).

Outcomes of low importance and specific adverse events are not included.

**Table 2**

Recommended outcome categories in trials for abnormal uterine bleeding

Outcome category	Key symptoms or outcomes
Bleeding	<ul style="list-style-type: none"> <li>• Amount</li> <li>• Frequency</li> <li>• Duration</li> <li>• Irregularity and pattern</li> <li>• Symptom bother</li> <li>• Primary problem and if resolved after treatment</li> </ul>
Quality of life	<ul style="list-style-type: none"> <li>• Disease-specific measures preferable</li> <li>• Generic measures can augment bleeding-specific measures to make comparison with other disease processes</li> </ul>
Pain	<ul style="list-style-type: none"> <li>• Bleeding-related pain</li> </ul>
Sexual health	<ul style="list-style-type: none"> <li>• Disease-specific measures preferable</li> <li>• Generic sexual health measures can augment bleeding-specific measures to make comparisons with other disease processes</li> </ul>
Bulk-related symptoms	<ul style="list-style-type: none"> <li>• For studies in women with uterine leiomyomas</li> </ul>
Patient satisfaction	<ul style="list-style-type: none"> <li>• Self-explanatory</li> </ul>
Additional treatments	<ul style="list-style-type: none"> <li>• Additional conservative or surgical treatments required during minimum 1-yr recommended follow-up period</li> </ul>
Adverse events	<ul style="list-style-type: none"> <li>• Clearly define adverse events</li> <li>• Seriousness and severity of event</li> <li>• Expected or unexpected event</li> <li>• Specify number and type of each type of adverse event in each study arm</li> <li>• Describe the number and type of any procedures or hospitalizations related to the first intervention within 30 d of intervention or initiating therapy per arm</li> </ul>