




## RESEARCH ARTICLE

# Standardising outcome reporting for clinical trials of interventions for heavy menstrual bleeding: Development of a core outcome set

N. A. M. Cooper<sup>1</sup>  | C. Rivas<sup>2</sup> | M. G. Munro<sup>3</sup> | H. O. D. Critchley<sup>4</sup> | T. J. Clark<sup>5</sup> | K. A. Matteson<sup>6</sup> | R. Papadantonaki<sup>7</sup> | S. Yorke<sup>1</sup> | A. Tan<sup>1</sup>  | M. Bofill Rodriguez<sup>8</sup> | M. Bongers<sup>9</sup> | A. Al-Hendy<sup>10</sup> | L. Bahamondes<sup>11</sup> | A. Connolly<sup>12</sup> | C. Farquhar<sup>8</sup> | T. Gray Valbrun<sup>13</sup> | M. Hickey<sup>14</sup> | H. S. Taylor<sup>15</sup> | D. Toub<sup>16</sup> | S. Vannuccini<sup>17</sup>  | S. Iliodromiti<sup>1</sup> | K. Khan<sup>18,19</sup>

<sup>1</sup>Women's Health Research Unit, Wolfson Institute of Population Health, Queen Mary University London, London, UK

<sup>2</sup>Social Research Institute, UCL Institute of Education, London, UK

<sup>3</sup>Department of Obstetrics & Gynecology, David Geffen School of Medicine, University of California, Los Angeles, California, USA

<sup>4</sup>MRC Centre for Reproductive Health, University of Edinburgh, Edinburgh, UK

<sup>5</sup>Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

<sup>6</sup>University of Massachusetts Chan Medical School, Worcester, Massachusetts, USA

<sup>7</sup>Maxima Medisch Centrum, Veldhoven, The Netherlands

<sup>8</sup>Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand

<sup>9</sup>Department of Obstetrics and Gynaecology, Grow-School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands

<sup>10</sup>Department of Obstetrics and Gynecology, University of Chicago, Chicago, Illinois, USA

<sup>11</sup>Department of Obstetrics and Gynaecology, School of Medical Sciences, Universidade Estadual de Campinas, Campinas, Brazil

<sup>12</sup>Bevan Healthcare, Bradford, UK

<sup>13</sup>The White Dress Project, Lithia Springs, Georgia, USA

<sup>14</sup>Department of Obstetrics and Gynaecology, University of Melbourne and The Royal Women's Hospital, Melbourne, Victoria, Australia

<sup>15</sup>Department of Obstetrics, Gynecology and Reproductive Sciences, Yale School of Medicine, New Haven, Connecticut, USA

<sup>16</sup>Medical Affairs, Gynesonics, Redwood City, California, USA

<sup>17</sup>Department of Obstetrics and Gynaecology, Careggi University Hospital, Florence, Italy

<sup>18</sup>Department of Preventive Medicine and Public Health, University of Granada, Faculty of Medicine, Granada, Spain

<sup>19</sup>CIBER Epidemiology and Public Health, Madrid, Spain

## Correspondence

N. A. M. Cooper, Women's Health Research Unit, Queen Mary University of London, Yvonne Carter Building, 58 Turner Street, London E1 2AB, UK.  
Email: [natalie.cooper@qmul.ac.uk](mailto:natalie.cooper@qmul.ac.uk)

## Funding information

Academy of Medical Sciences; National Institute for Health Research; Beatriz Galindo (senior modality) Programme Grant

## Abstract

**Objective:** To develop a core outcome set for heavy menstrual bleeding (HMB).

**Design:** Core outcome set (COS) development methodology described by the COMET initiative.

**Setting:** University hospital gynaecology department, online international survey and web-based international consensus meetings.

**Population or sample:** An international collaboration of stakeholders (clinicians, patients, academics, guideline developers) from 20 countries and 6 continents.

**Methods:** Phase 1: Systematic review of previously reported outcomes to identify potential core outcomes. Phase 2: Qualitative studies with patients to identify outcomes most important to them. Phase 3: Online two-round Delphi survey to achieve consensus about which outcomes are most important. Phase 4: A consensus meeting to finalise the COS.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *BJOG: An International Journal of Obstetrics and Gynaecology* published by John Wiley & Sons Ltd.

**Main outcome measures:** Outcome importance was assessed in the Delphi survey on a 9-point scale.

**Results:** From the 'long list' of 114, 10 outcomes were included in the final COS: subjective blood loss; flooding; menstrual cycle metrics; severity of dysmenorrhoea; number of days with dysmenorrhoea; quality of life; adverse events; patient satisfaction; number of patients going on to have further treatment for HMB and haemoglobin level.

**Conclusions:** The final COS includes variables that are feasible for use in clinical trials in all resource settings and apply to all known underlying causes of the symptom of HMB. These outcomes should be reported in all future trials of interventions, their systematic reviews, and clinical guidelines to underpin policy.

#### KEY WORDS

AUB, core outcome set, heavy menstrual bleeding

## 1 | INTRODUCTION

Heavy menstrual bleeding (HMB) is a debilitating condition with a significant impact on women's physical, psychological and social wellbeing<sup>1-5</sup> and an economic impact due to time taken away from employment and costs to healthcare systems.<sup>6-8</sup> The symptom of HMB may be ascribed to multiple underlying causes<sup>9</sup> and there are a variety of effective treatments, including hormonal, medical and surgical interventions.<sup>2,5</sup> Although these treatments have been widely explored in clinical trials, many different outcome measures have been used. This heterogeneity has limited treatment comparisons and comprehensive data syntheses, thus diluting the strength of recommendations made by clinical guidelines.<sup>1,2</sup> Clear definitions and standardised reporting are needed to maximise the interpretability of clinical research and enable a better understanding of treatment effectiveness in HMB.

Although checklists exist for the reporting of clinical trials, core outcome sets (COS) differ because they are a disease-specific agreed set of outcomes that are established as a reporting standard minimum for all relevant clinical trials. The aim of COS is to ensure that studies of a condition all report the same, valid outcomes, which will ultimately mean they will produce results that are not only useful for interpretation of that trial but can also contribute to meta-analyses and the overall assessment of interventions.

There is an urgent need for a COS for HMB which as well as addressing the aforementioned issues, would improve the quality of trial reporting, prevent selective reporting and reduce research waste. These standardised outcomes should be considered key metrics when developing clinical guidance and health care policy.

We, therefore, developed a COS to be used in all future trials evaluating interventions for treating HMB regardless of the underlying cause or the type of intervention.

## 2 | METHODS

The COS was developed using methods described by the COMET (Core Outcome Measures in Effectiveness Trials) Initiative<sup>10</sup> and pre-registered on their database. The project had four distinct phases.

### 2.1 | Phase 1: Systematic review of previously reported outcomes in studies examining interventions for HMB

A systematic review of previously reported outcomes in studies of HMB symptoms was conducted and has been reported elsewhere.<sup>11</sup> In brief, medical databases and trial registries were searched, and systematic reviews were cross-referenced to identify randomised controlled trials (RCTs) and observational studies that explored interventions for HMB. All primary and secondary outcomes were extracted and used to develop a 'long list' to be considered for the COS.

### 2.2 | Phase 2: Qualitative studies with patients

The 'long list' derived from Phase 1 was used in the design of a workshop and interviews involving patients and their partners to obtain qualitative data about which outcomes were perceived to be most important. Participants were recruited from a teaching hospital gynaecology clinic. They were eligible if they were over 18 years old, if they or their partner had a history of HMB, and if they had a reasonable understanding of written and spoken English. Data were analysed to identify additional outcomes to add to the 'long list'. The transcripts were used to inform latter phases of the project, ensuring that patients were given a voice. The workshop methodology was shared with other researchers to allow them to conduct their own qualitative projects, increasing

the pool of patients contributing and the reach of the findings. Researchers in Chile and the Netherlands contributed to the qualitative work in this way. Full details of phase 2 have been published elsewhere.<sup>12</sup>

### 2.3 | Phase 3: Online Delphi survey

Using the 'long list' of outcomes, a two-round Delphi survey was conducted. The Delphi technique is a systematic process for developing and measuring consensus.<sup>13,14</sup> DELPHI MANAGER software<sup>15</sup> was used to create and manage the survey. We planned to conduct the Delphi survey in three rounds but, during the course of the project, COMET recommended that two rounds should be sufficient and due to time limitations posed by the COVID-19 pandemic, restricting the number of rounds was considered expedient.

The following stakeholders were invited via email; major national and international gynaecological societies, editors of gynaecological journals, participants from the qualitative workshops, health technology companies and HMB guideline developers. We used a 'snowball approach', asking recipients to forward our invitation to their contacts with an interest in HMB. To identify additional patients from across the globe, we collaborated with the Clue® (BioWink GmbH) menstrual tracking app, advertising the survey via their English-speaking newsletter.

Round 1 of the survey was open from 22 January to 21 March 2020. Stakeholders were divided into three groups: (i) clinician (doctor, nurse, allied health professional, other healthcare provider, researcher/academic, journal editor); (ii) guideline developer, policy maker, service commissioner, other; and (iii) patient, patient's partner or family member. The 'long list' of outcomes was presented in subgroups, and participants were asked to rate the importance of the outcomes on a scale of 1–9, with 1–3 labelled 'not important', 4–6 labelled 'important but not critical', and 7–9 labelled 'critical.' Each outcome was presented with a 'lay description' written following consultation with a patient and public involvement (PPI) group (Katie's team <https://www.barc-research.org/katies-team>)<sup>16</sup> connected to the research areas of childbirth, pregnancy and reproductive issues. At the end of the survey, participants were asked to suggest additional outcomes.

Participants who had completed Round 1 were invited to Round 2, which ran between 31 March 2020 and 31 May 2020. The list of outcomes was presented again but participants were shown how they had individually scored the outcome in the first round and how each stakeholder group had rated it. This allowed them to consider each group's score before re-scoring the outcome and led to a convergence in thinking. No outcomes were removed from the survey between the rounds.

The criteria for determining which outcomes should be included in the COS (consensus in) required more than 70% of the participants rating an outcome as 'critical' while fewer than 15% rated it as 'not important'. The converse

determined 'consensus out' outcomes. We planned to use the RAND Disagreement index<sup>17</sup> if consensus could not be reached about outcomes. If consensus was still not reached for an outcome, further evaluation was required in the final phase of the project's consensus meeting.

### 2.4 | Phase 4: Consensus meeting

Initial plans to hold the consensus meeting at an international conference had to be adapted due to the COVID-19 pandemic. Instead, we arranged a video teleconference meeting with stakeholders' representatives to discuss the potential core outcomes and identify which should become the final COS.

Delphi participants who were members of the FIGO Committee on Menstrual Disorders and the Society for Endometriosis and Uterine Disorders (SEUD) Abnormal Uterine Bleeding Task Force were invited to represent the global clinical and academic community. One of these clinicians was also a medical device company physician and medical director. A health economist with guideline development experience and a general practitioner with commissioning experience were also invited to ensure that the COS would be applicable across healthcare systems and relevant to guideline development. Patients from phases 2 and 3 were invited to attend, as was the founder of a patient advocacy group, The White Dress Project, who had experience in advocating for women in groups of clinical professionals.

The number of outcomes remaining after round 2 was prohibitively high and was not reduced by RAND analysis; thus, prior to the first meeting, we agreed with the consensus group participants that if any subgroup (clinicians, patients, others) had more than 50% of participants who thought that the outcome was critical (rated 7–9), then it remained in for discussion. This gave all of the subgroups an equal say.

The meeting was arranged for a mutually convenient time; to accommodate multiple time zones and account for attendees' ability to engage actively in productive discussion, we decided to schedule two separate meetings, each lasting 2 hours.

We used Blackboard Collaborate (Blackboard Inc.) to conduct an interactive meeting with video conferencing facilitated by two authors (NAM and CR). NAMC introduced the meeting and the format and explained the graphical depictions of the data. CR then facilitated discussions, particularly ensuring that the views of the patient representative were considered, using findings from the qualitative work (phase 2) to support their perspective. The 'consensus in' outcomes and the 'no consensus' outcomes were discussed. The aim was to negotiate a shared understanding that enabled consensus. Often, discussions made it clear whether an outcome should be included or not but participants were asked formally to vote on whether an outcome should be excluded or included. An agreement was achieved at a level of 70%. Some outcomes were re-evaluated if a later-discussed

outcome appeared to evaluate the same endpoint (e.g. measurement tool assessment of menstrual blood loss versus a subjective evaluation of menstrual blood loss), meaning that we looked at the outcomes as relevant to one another rather than individually.

The patient advocate and the primary care/commissioning representative were not able to join the second meeting due to last minute, unforeseen circumstances. It was decided to conduct the planned consensus meeting but to discuss the findings with these participants before making the final decisions regarding which outcomes to include. The final COS was then confirmed with all participants via group email communications.

After the meeting, outcomes rated as 'consensus in' by the Delphi survey but not included in the final core outcome set were reviewed and mapped to the final core outcome set to identify any aspects that might need to be re-evaluated.

### 3 | RESULTS

#### 3.1 | Phases 1 and 2: Systematic review and qualitative studies with patients

The systematic review<sup>11</sup> identified 136 outcomes which were consolidated to a list of 109 for the Delphi survey (Appendix S1). An additional five outcomes were identified during the qualitative work and added to the survey<sup>12</sup> (see Table S1).

#### 3.2 | Phase 3: Delphi survey

Round 1 of the Delphi survey contained 114 outcomes. Thirty additional outcomes were suggested in Round 1. Of these suggestions, 19 were excluded (already included  $n=2$ ; not 'outcomes'  $n=11$ ; duplicates  $n=5$ ; specific to fibroids  $n=1$ ) and 11 outcomes were added to the survey for Round 2. No outcomes were removed between rounds (see Table S1).

Participants in the Delphi survey and consensus meetings reflected all intended stakeholders. The majority of participants were clinicians. Although patients were invited to the final consensus meeting, only one patient representative agreed to join. Table S2 shows the characteristics of participants in the development of the core outcome set (COS).

After analysis, 38 outcomes met 'consensus in' criteria, none met 'consensus out' criteria and 87 had no consensus. Using the RAND Disagreement Index only allowed us to remove one 'no consensus' outcome, so we applied agreed additional criteria and 25 'no-consensus' outcomes were removed. This left 100 outcomes (all 38 'consensus in' outcomes and 62 'no consensus' outcomes) for discussion. Table S1 lists the outcomes considered at each stage of the consensus process.

### 3.3 | Phase 4: Consensus meeting

The final COS is shown in Table 1.

When reviewing 'consensus in' outcomes that were not included in the final COS, it was identified that many were encompassed by broader outcomes included in the COS (see Table 2).

## 4 | DISCUSSION

### 4.1 | Main findings

This COS was developed using rigorous, standardised methodology<sup>10</sup> involving a global group of participants from 20 countries and 6 continents, representing all major stakeholder groups. The COS was finalised through deliberative discussion that upheld the patient's voice and practical clinical needs. It included variables that are feasible for use in clinical trials in all resource settings and apply to all known underlying causes of the symptom of HMB.

There were multiple outcomes included in the Delphi survey which centred on subjective assessment of menstrual blood loss. Although we did not intend to specify a measure, we felt it was important to keep these measurements in the survey as distinct outcomes, as they are commonly the primary outcome in a study, and there are so many diverse measures. Just having 'menstrual blood loss' as an outcome would not prevent various measures from being used, and this would not achieve our aim of making study outcomes more homogeneous. Many of the menstrual blood loss (MBL) outcomes met 'consensus in' criteria after Phase 3. The consensus group identified that for this outcome to be relevant to all potential interventions, it should allow for recording of the full range of menstrual blood loss options. Using amenorrhoea as a binary outcome has utility limited to a small subset of interventions (e.g. hysterectomy, endometrial ablation). Our patient participants stressed that numeric representation of MBL or absence of bleeding was not the outcome most meaningful for them; they preferred treatments to make their bleeding 'better'. Some regulatory bodies (e.g. U.S. Food and Drug Registration [FDA]), mandate the use of quantitative or semiquantitative measures to show improvement in MBL before approving new interventions, but as these numbers mean little to patients, or even to many gynaecologists, it seems more appropriate to apply assessments that reflect real-life benefits to women. Additionally, using alkaline haematin testing or PBLAC scores is labour-intensive, unappealing to women and costly.

Flooding is a term that is not universally recognised, describing sudden overwhelming blood loss that exceeds the saturation of the menstrual products being used (often soaking through clothing and on to furnishings). This is a symptom that stops women leaving the house, restricts their daily lives and frequently causes them embarrassment. Our patient participants expressed how the unpredictability of bleeding is the most challenging thing to tackle. Describing

**TABLE 1** The core outcome set for the symptom of heavy menstrual bleeding.

Outcome	Comments
Subjective blood loss	In the Delphi survey, all of the subjective techniques (VAS, categories, i.e. light through to heavy or worse through to better, Likert scales, etc.) were considered 'consensus in' outcomes and they were all rated more important than the quantitative ones (alkaline haematin, Pictorial blood loss assessment chart [PBLAC]). The highest rated outcome within the survey was 'Subjective assessment of change in menstrual bleeding from baseline (e.g. greatly improved to much worse)' with 95% of participants stating it was 'critical' and 0% saying it was 'not important'
Flooding <sup>a</sup>	This outcome did not meet the 'consensus in' criteria after the Delphi survey despite 87% of patients and 80% of other participants rating it as critical. At the consensus meeting it was discussed as a 'no consensus' outcome and met criteria to become a core outcome. We recognise that a formal definition of flooding needs development
Menstrual cycle metrics	None of the cycle metrics met the Delphi 'consensus in' criteria but were discussed in the consensus group meeting as 'no consensus' outcomes. During discussions, they were highlighted as key factors for assessing 'normal' versus 'abnormal' menstruation and met the criteria for core outcomes
Severity of dysmenorrhoea Number of days with dysmenorrhoea	These were both 'consensus in' outcomes after the Delphi survey. The consensus group discussed the appropriateness of dysmenorrhoea in the COS for HMB and considered the option of such outcomes being part of a dedicated COS for pelvic pain. However, in our qualitative work, many patients stressed the importance of pain on their quality of life
Condition-specific quality of life measure and a generic quality of life measure	Both generic and condition-specific quality of life were considered 'consensus in' outcomes after the Delphi survey. The consensus group agreed that they should form part of the COS as the impact of menstrual bleeding and menstruation symptoms on quality of life is at the core of the definition of HMB, as specified by the NHS NICE, FIGO and the American College of Obstetrics and Gynecology (ACOG) <sup>1,9,27,28</sup>
Adverse events and any relevant prespecified safety measures	All groups consistently rated adverse events and treatment-specific safety measures as critical in the Delphi survey to become 'consensus in' outcomes. The consensus group agreed that they should be included within the COS, in keeping with principles of Good Clinical Practice. Treatment-specific safety measures were recognised as falling under the same category but requiring prespecification dependent upon the evaluated interventions, e.g. blood pressure monitoring with combined estrogen- and progestin-containing contraceptives
Patient satisfaction	This was a 'consensus in' outcome after the Delphi survey, and in the consensus meeting, participants identified that the expressed satisfaction of the patient with the intervention is an important assessment of treatment because this global measure addresses the whole experience, not just a single aspect.
Number of patients going on to have further treatment for HMB	In the Delphi survey, women rated this outcome very highly and it was a 'consensus in' outcome. It is an aspect of effectiveness and an outcome important for patients who want to know whether interventions can be expected to have 'longevity'
Haemoglobin	Haemoglobin, haematocrit, ferritin levels and other iron parameters were initially 'no consensus' outcomes. However, iron deficiency is a common consequence of the symptom of HMB which, even in the absence of anaemia, may adversely impact cognitive and physical function. The consensus group decided that it was essential to include one of these parameters of iron deficiency and thus haemoglobin became a 'consensus in' outcome

<sup>a</sup>Sudden overwhelming blood loss that exceeds the saturation of the menstrual products being used.

cycle metrics (e.g. using the FIGO AUB System 1)<sup>9</sup> will give women information about what to expect after treatment and is important for treatments that cause unscheduled bleeding as side effects.

Validated measures of health-related quality-of-life are essential to provide data regarding a change in the impact of menstrual symptoms on how a woman feels and functions. Generic quality-of-life measurement tools allow the comparison of different types of medical conditions and symptoms. However, as highlighted by discussions in the consensus meeting, HMB is a cyclical symptom, so a generic measure will not necessarily pick up adverse effects on quality of life if it is assessed on a non-menstruating day. Therefore, a condition-specific validated patient-reported outcome measure (PRO) can accurately evaluate the quality of life of women with HMB. Patient satisfaction can be used to assess easily the different types of treatment. Patients have different expectations and thus satisfaction evaluates how well those expectations are met.

A challenge of 'retreatment' as an outcome is defining the time frame for data capture. In the systematic review of

outcomes,<sup>11</sup> the most common time point for reporting results was at 1 year, but other frequently reported milestones were 3, 6 and 24 months.

Iron deficiency (ID) is a common consequence of the symptom of HMB. Iron deficiency (ID), with or without iron deficiency anaemia (IDA), may compromise pregnancy and neonatal outcomes, including fetal neurodevelopment.<sup>18,19</sup> Also, morbidity and mortality are increased when women undergoing gynaecological surgery are anaemic.<sup>20</sup> Post-intervention iron supplementation is important, as evidence demonstrates that years after intervention, women can still have ID.<sup>21</sup> Consequently, managing those with HMB should include assessment and treatment of ID and IDA. The consensus group considered haemoglobin, haematocrit, ferritin levels and other iron parameters. However, in many environments, especially those with a high prevalence of coexisting inflammatory conditions, ferritin levels may be deceptively high and ferritin measurement is also a relatively expensive test. Consequently, including ferritin may have been prohibitive in low-resource settings. Based on these factors, haemoglobin became a 'consensus in' outcome.

**TABLE 2** How the 'consensus in' criteria that were not included in the final core outcome set for heavy menstrual bleeding are encompassed by individual core outcomes.

Core outcome	The 'consensus in' outcomes that were not included in the core outcome but are encompassed by the adjacent core outcome
Subjective blood loss	Severity of menstrual blood loss reported with a numerical score to represent severity Severity of menstrual blood loss reported in groups of increasing volume Subjective assessment of change in menstrual bleeding from baseline Number of women whose symptoms have not changed or have become worse Number of days of heavy bleeding
Flooding	Number of women passing blood clots during menstruation Number of days of heavy bleeding
Menstrual cycle metrics	Number of days of heavy bleeding
Severity of period pain	Severity of pain
Number of days with period pain	–
Condition-specific quality of life measure and generic quality of life measure	Effect of menstruation on professional activities or education Effect of menstruation on activities of everyday living Number of days a woman is unable to leave the house due to menstruation Duration of time away from professional activities or education due to menstruation Time to resumption of normal activities
Adverse events and any relevant prespecified safety measures	Compliance with treatment Treatment acceptability at follow-up Severity of postoperative pain at 2–14 days
Patient satisfaction	Patient's assessment of effectiveness of treatment Successful treatment Discontinuation with treatment Psychological wellbeing Compliance with treatment Cost to patient
Number of patients undergoing retreatment for HMB	Time to re-intervention How long it takes for periods to become heavy again after treatment stops Number of failed procedures
Haemoglobin	–
<b>Consensus in outcomes that were not linked to a core outcome</b>	
<b>Outcome</b>	<b>Comment</b>
Time to full recovery	Excluded as hard to define 'full recovery' and relevant to a woman's baseline status
Length of stay in the hospital/clinic/medical centre	Excluded, as this is driven by many things such as type of anaesthesia, type of healthcare system, patient choice
Number of completed procedures	Excluded, as only specific to procedural interventions and it evaluates feasibility rather than the effectiveness of an intervention

Despite not all of the 'consensus in' outcomes making it through to the final COS, we identified that most of those not included directly were encompassed by outcomes that did form the COS. Thus the final set of outcomes is an accurate representation of our stakeholders' opinions. The three outcomes that were not represented were difficult to define or were not applicable or comparable across interventions and healthcare settings.

## 4.2 | Strengths and limitations

Developing the COS for HMB faced challenges. The project was extended due to the primary researcher taking parental leave and the need to update the initial outcome review prior to undertaking qualitative studies. Although over 40 patients were recruited for our qualitative workshop, only six

attended; however, these women were representative of different age groups and ethnicities, and one male participant was interviewed. We supplemented phase 2 with telephone interviews and findings from workshops in the Netherlands and Chile.<sup>12</sup> We had high attrition rates in the patient group in the Delphi survey (phase 3).

We were able to involve participants from across the world and a variety of resource settings. Ironically, this degree of participation was likely aided by the COVID-19 pandemic by necessitating a change from a planned face-to-face consensus meeting to a web-based video conference.

## 4.3 | Interpretation

Previous work examining outcomes in randomised controlled trials for HMB identified a lack of consistency and

the need for standardised reporting.<sup>22,23</sup> This project has built on that work by evaluating all outcomes across all study types and by working with stakeholders to develop the COS and standardise future reporting. We have not yet explored which measurement tools to use. We recognise the frustration that this generates for researchers trying to use the COS for their work. Strict methodology exists for reviewing outcome measures;<sup>24</sup> however, the criteria for tools to be acceptable are difficult to satisfy and risk no recommendation being made. We plan to conduct further work to identify appropriate tools for measuring the specified outcomes in the COS with pragmatic results providing a recommendation for each one.

#### 4.4 | Practical implications

It is important to acknowledge that the COS reflects a minimum reporting dataset for studies on HMB rather than an exhaustive set of outcomes. Investigators should report the full COS but are welcome to report as many additional outcomes if they wish. If a COS outcome cannot be used, the investigators should justify this.

Core outcomes should be feasible for use in all research settings including LMIC; thus, simple outcome measures that are freely available and easy to use are likely to be most appropriate. In addition, specifying follow-up time points would further standardise the construct of outcome data.

There are 10 outcomes in the COS; however, outcome tools may be identified that assess more than a single core outcome. For example, the Menstrual Bleeding Questionnaire (MBQ)<sup>25</sup> comprises items on menstrual bleeding, pain and quality of life (QoL), with a focus on social embarrassment and alterations to daily activities specific to the symptom of HMB. Combining the number of days and severity of dysmenorrhoea would allow assessment of day-to-day severity, thus creating an index better expressed as an 'area under the curve'. Further work to specify outcome assessment tools is likely to reduce the amount of work needed to assess and report the COS.

#### 4.5 | Research implications

Development of this COS has identified a need to standardise aspects of clinical investigation of interventions for HMB. Inherently, the goal of COS design and implementation is to encourage standardised reporting of relevant outcomes. It is important to understand that HMB is a symptom many women offer as their chief complaint, not a diagnosis, and for data to be compared or synthesised, populations should be homogeneous concerning the underlying cause. With this in mind, we encourage the use of FIGO's AUB System 2, the PALM-COEIN classification<sup>9</sup> to describe baseline populations in studies of HMB. Research has identified the increasing use of patient-reported outcome measures (PROs)<sup>26</sup> and the need for validated outcome tools for accurate assessment

of treatments for HMB in a way that represents the experiences of those suffering from the symptom.

#### AUTHOR CONTRIBUTIONS

NAMC, CR and KSK developed the methodology and secured funding and ethical approval. MGM, HODC, TJC, RP, MBR and MB also contributed to design aspects of the project. NAMC, CR, RP, SY, AT, MBR and MB were involved in acquisition of data. All authors were involved in analysis and interpretation of the data and drafting or revision of the manuscript. All authors approved the final submitted version of the paper and have agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

#### ACKNOWLEDGEMENTS

With thanks to Professor Steph Taylor (Public Health and Primary Care, Queen Mary University of London, London, UK) and Professor Peter O'Donovan (University of Bradford, Bradford, UK) who advised during setup and Phase 1 of the project; Professor Paula Williamson from the COMET initiative and University of Liverpool, who advised on methodology; Dr Johana Nayoan, at the time a Research Fellow for Blue Communities Research Programme, University of Exeter, who facilitated the patient workshop; Katie's Team Patient and Public Involvement group ([www.barc-research.org/katies-team](http://www.barc-research.org/katies-team)), who assisted with development of our patient documents and the Delphi survey; Dr Javier Zamora (Clinical Biostatistics Unit, Hospital Ramón y Cajal, Madrid, Spain) for statistical support; The Clue® (BioWink GmbH, Berlin, Germany) menstrual tracking app for promoting our Delphi survey to their members; Dr Paul Jacklin (National Institute for Health and Care Excellence, UK), Dr Margit Duelhom (Aarhus Universitet Health, Denmark), Dr Ilza Monteiro (Department of Obstetrics and Gynaecology, Universidade Estadual de Campinas, Brazil), Prof. Evan Myers (Department of Obstetrics and Gynecology, The University of North Carolina at Chapel Hill, USA), Prof. John Thiel (College of Medicine, University of Saskatchewan, Canada) and Prof. Ricardo Lasmar (Universidade Federal Fluminense, Brazil) for participating in the consensus meetings; Prof. Jason Abbott (School of Clinical Medicine, Health and Medicine Division of Obstetrics and Gynaecology, University of New South Wales, Australia), Dr Delfin A. Tan (Department of Obstetrics and Gynaecology, St. Luke's Medical Center, Philippines), Dr Ally Murji (Department of Obstetrics and Gynaecology, University of Toronto, Canada) and Dr Sarah Hillman (Primary Care, University of Warwick, UK) for advice on the final manuscript.

#### FUNDING INFORMATION

The core outcomes set for HMB was funded by an Academy of Medical Sciences Starter Grant for Clinical Lecturers which was awarded to Dr Cooper, who during this project was an Academic Clinical Lecturer funded by the National Institute for Health Research. Professor Khan is a

Distinguished Investigator funded by the Beatriz Galindo (senior modality) Programme Grant given to the University of Granada by the Spanish Ministry of Education.

### CONFLICT OF INTEREST STATEMENT

NAMC, TJC, RP, SY, AT, MBR, LB, MH and DBT have nothing to disclose. CR reports grants paid to her institution from the National Institute for Health Research, UK Research and Innovation and University College London; consulting fees for OFFICE FOR STUDENTS; support for attending meetings from COST European consortium; participation in advisory/data monitoring boards for Vaccines Lambeth and Autism and HE; and leadership roles with National Institute for Health and Health Care Research (NIHR) Health and Social Care Delivery Research panel, Autistica and the Ehlers Danlos society. MGM reports receiving consulting fees from American Regent, AbbVie Inc., Caldera Medical, Channel Medical, Clue, CooperSurgical, Daiichi-Sankyo, Myovant Sciences, Pharmacosmos, Okon Medical and Rejoni Inc.; payment as a speaker from Daiichi-Sankyo and the Women's Health Research Collaborative; support for attending meetings from Vifor Medical; stock options from Channel Medical, Gynesonics Inc., CrossBay Medical, Okon Medical and Rejoni Inc.; and leadership roles in the Women's Health Research Collaborative, FIGO Menstrual Disorders and Related Health Impacts and SEUD AUB Task Force. HODC reports grants paid to her institution from the Medical Research Council, Bayer AG, NIHR and the Biotechnology and Biological Sciences Research Council; consulting fees paid to her institution from Myovant Sciences GmbH, Vifor Pharma UK Ltd, Gedeon Richter, Bayer AG; royalties from Up-to-Date; support for attending meetings from Myovant Sciences GmbH and Bayer AG; and, participation in advisory/data monitoring boards for Myovant Sciences GmbH, Bayer AG and Wellbeing of Women RAC. KAM reports consulting fees received from Myovant Sciences and Bayer AG and a leadership role with the American College of Obstetrics and Gynecology. MB reports payment from Hologic Inc. for lecture fees; participation on the advisory board of Gedeon Richter; and is Chair of the Dutch interest group of benign gynaecology. AA-H reports grants from the National Institutes of Health; consulting fees from AbbVie, Bayer AG, ObsEva and Myovant Sciences; and a patent for novel diagnostics and therapeutics for uterine sarcoma. AC reports payment for lectures given from Bayer AG, Gedeon-Richter and Hologic Inc. CF reports an advisory role on the TILT (Time-Lapse Imaging Trial) randomised controlled trial; and a leadership role as dean of policy and research for the Royal Australian and New Zealand College of Obstetricians and Gynaecologists (RANZCOG). HST reports a grant award from AbbVie; a leadership role with the American Society for Reproductive Medicine; and receipt of medication and its placebo from AbbVie. SV reports consulting fees and payments for lectures from Bayer AG. TGV reports grants to The White Dress Project from Myovant Sciences GmbH, Hologic Inc. and Pfizer. She has received travel grants and honoraria from Myovant Sciences GmbH and Gynesonics Inc. She is the founder of The White Dress

Project. SI reports grants from the Health Foundation and The Healthcare Improvement Studies (THIS) Institute. KSK reports receiving fees for giving expert testimony in legal cases; for receiving fees and expenses for giving lectures for Ferring Pharmaceuticals, Olympus and various Universities and Societies; receiving royalties from Taylor and Francis, Hodder Arnold, and Huber, for books authored by him; is a partner/director in a Partnership/Company (HER [Health Education Research Ltd, UK]) where he offers advice, including matters within medical research, manuscripts and negligence; and sponsorship for meetings from Ferring Pharmaceuticals, Leo-Pharma, Alere, Ethicon, Hologic Inc., Viforpharma and Preglem/Quintiles. Completed disclosure of interest forms are available to view online as supporting information.

### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### ETHICS APPROVAL

Ethical approval for this project was granted by the East Midlands – Nottingham 1 Research Ethics Committee on 14 December 2015, REC reference ID 15/EM/0565. All patients involved in the qualitative workshops and interviews provided written consent for the storage and use of their contact details and participation in the project. All Delphi survey participants gave their consent to participate when registering for the survey.

### ORCID

N. A. M. Cooper  <https://orcid.org/0000-0003-1929-833X>

A. Tan  <https://orcid.org/0000-0002-4842-9987>

S. Vannuccini  <https://orcid.org/0000-0001-5790-587X>

### REFERENCES

1. National Institute for Health and Care Excellence. Heavy menstrual bleeding: assessment and management. 2018 [cited 2 September 2022]. Available from: <https://www.nice.org.uk/guidance/cg44>
2. National Institute for Health and Care Excellence. Heavy menstrual bleeding. London: RCOG Press; 2007.
3. Coulter A, Peto V, Jenkinson C. Quality of life and patient satisfaction following treatment for menorrhagia. *Fam Pract*. 1994;11(4):394–401.
4. Jones GL, Kennedy SH, Jenkinson C. Health-related quality of life measurement in women with common benign gynecologic conditions: a systematic review. *Am J Obstet Gynecol*. 2002;187(2):501–11.
5. 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(2):96–104.
6. 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(2):343–8.
7. Frick KD, Clark MA, Steinwachs DM, Langenberg P, Stovall D, Munro MG, et al. Financial and quality-of-life burden of dysfunctional uterine bleeding among women agreeing to obtain surgical treatment. *Womens Health Issues*. 2009;19(1):70–8.
8. Schoep ME, Adang EMM, Maas JWM, De Bie B, Aarts JWM, Nieboer TE. Productivity loss due to menstruation-related symptoms: a nationwide cross-sectional survey among 32 748 women. *BMJ Open*. 2019;9(6):e026186.
9. Munro MG, Critchley HOD, Fraser IS, FIGO Menstrual Disorders Committee. The two FIGO systems for normal and abnormal uterine



- bleeding symptoms and classification of causes of abnormal uterine bleeding in the reproductive years: 2018 revisions. *Int J Gynecol Obstet*. 2018;143:393–408.
10. COMET Initiative. Core outcome measures in effectiveness trials. 2014 [cited 2 September 2022]. Available from: <http://www.comet-initiative.org>
  11. Cooper NAM, Yorke S, Papadantonaki R, Khan KS. Variation of outcome reporting in studies of interventions for heavy menstrual bleeding: a systematic review. *Facts Views Vis Obgyn*. 2022;14(3):205–18. <https://doi.org/10.52054/FVVO.14.3.030>
  12. Cooper NAM, Yorke S, Tan A, Khan KS, Rivas C. A qualitative study exploring the effect of heavy menstrual bleeding on women and their close associates: stakeholder involvement in development of a core outcome set. *BMJ Open*. 2023, in press.
  13. Humphrey-Murto S, Wood TJ, Gonsalves C, Mascioli K, Varpio L. The Delphi method. *Acad Med*. 2020;95(1):168.
  14. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ*. 1995;311(7001):376–80.
  15. COMET initiative. DelphiManager [cited 2 September 2022]. Available from: <http://www.comet-initiative.org/delphimanager/>
  16. Moss N, Daru J, Lanz D, Thangaratnam S, Khan KS. Involving pregnant women, mothers and members of the public to improve the quality of women's health research. *BJOG*. 2017;124(3):362–5.
  17. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lázaro P, et al. The RAND/UCLA appropriateness method user's manual. Santa Monica: RAND; 2001 [cited 12 January 2022]. Available from: [https://www.rand.org/content/dam/rand/pubs/monograph\\_reports/2011/MR1269.pdf](https://www.rand.org/content/dam/rand/pubs/monograph_reports/2011/MR1269.pdf)
  18. Pratt JJ, Khan KS. Non-anaemic iron deficiency – a disease looking for recognition of diagnosis: a systematic review. *Eur J Haematol*. 2016;96(6):618–28.
  19. Georgieff MK. Iron deficiency in pregnancy. *Am J Obstet Gynecol*. 2020;223(4):516–24.
  20. Richards T, Musallam KM, Nassif J, Ghazeeri G, Seoud M, Gurusamy KS, et al. Impact of preoperative anaemia and blood transfusion on postoperative outcomes in gynaecological surgery. *PLoS One*. 2015;10(7):e0130861.
  21. Peuranpää P, Heliövaara-Peippo S, Fraser I, Paavonen J, Hurskainen R. Effects of anemia and iron deficiency on quality of life in women with heavy menstrual bleeding. *Acta Obstet Gynecol Scand*. 2014;93(7):654–60.
  22. Herman MC, Penninx JPM, Geomini PM, Mol BW, Bongers MY. Choice of primary outcomes evaluating treatment for heavy menstrual bleeding. *BJOG*. 2016;123:1593–1598.
  23. Rahn DD, Abed H, Sung VW, Matteson KA, Rogers RG, Morrill MY, et al. Systematic review highlights difficulty interpreting diverse clinical outcomes in abnormal uterine bleeding trials. *J Clin Epidemiol*. 2011;64(3):293–300.
  24. CONsensus-based Standards for the selection of health Measurement INstruments (COSMIN) [cited 2 September 2022]. Available from: <https://www.cosmin.nl/>
  25. Matteson KA, Scott DM, Raker CA, Clark MA. The menstrual bleeding questionnaire: development and validation of a comprehensive patient-reported outcome instrument for heavy menstrual bleeding. *BJOG*. 2015;122(5):681–9.
  26. Matteson KA. Menstrual questionnaires for clinical and research use. *Best Pract Res Clin Obstet Gynaecol*. 2017;40:44–54.
  27. ACOG reVITALize gynecology data definitions website [cited 2 September 2022]. Available from: <https://www.acog.org/practice-management/health-it-and-clinical-informatics/revitalize-gynecology-data-definitions>
  28. Sharp HT, Johnson JV, Lemieux LA, Currihan C. Executive summary of the reVITALize initiative: standardizing gynecologic data definitions. *Obstet Gynecol*. 2017;129(4):603–7.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Cooper NAM, Rivas C, Munro MG, Critchley HOD, Clark TJ, Matteson KA, et al. Standardising outcome reporting for clinical trials of interventions for heavy menstrual bleeding: Development of a core outcome set. *BJOG*. 2023;00:1–9. <https://doi.org/10.1111/1471-0528.17473>