

## Consensus recommendations for measuring the impact of contraception on the menstrual cycle in contraceptive clinical trials

Amelia C. L. Mackenzie<sup>1</sup>, Stephanie Chung<sup>2,3</sup>, Emily Hoppes<sup>2</sup>, Nora Miller<sup>4</sup>, Anne E. Burke<sup>5</sup>, \* Sharon L. Achilles<sup>6</sup>, C. Leigh Allen<sup>7</sup>, Luis Bahamondes<sup>8</sup>, Diana L. Blithe<sup>7</sup>, Vivian Brache<sup>9</sup>, Rebecca L. Callahan<sup>2</sup>, Alice F. Cartwright<sup>2,a</sup>, Kathryn B. H. Clancy<sup>10</sup>, Enrico Colli<sup>11</sup>, Amanda Cordova-Gomez<sup>12</sup>, Elizabeth E. Costenbader<sup>2</sup>, Mitchell D. Creinin<sup>13</sup>, Hilary O. D. Critchley<sup>14</sup>, Gustavo F. Doncel<sup>15</sup>, Laneta J. Dorflinger<sup>2</sup>, Alison Edelman<sup>16</sup>, Thomas Faustmann<sup>17,b</sup>, Christoph Gerlinger<sup>17,18</sup>, Lisa B. Haddad<sup>19</sup>, Julie Hennegan<sup>20</sup>, Cássia Raquel Teatin Juliato<sup>8</sup>, Simon P. S. Kibira<sup>21</sup>, Diana Mansour<sup>22</sup>, Andres Martinez<sup>23</sup>, Kristen A. Matteson<sup>24</sup>, Jacqueline A. Maybin<sup>14</sup>, Alexandria K. Mickler<sup>12,c</sup>, Kavita Nanda<sup>2</sup>, Chukwuemeka E. Nwachukwu<sup>12</sup>, Funmilola M. OlaOlorun<sup>25</sup>, Kevin J. Peine<sup>12</sup>, Chelsea B. Polis<sup>19</sup>, Carolina Sales Vieira<sup>26</sup>, Regine Sitruk-Ware<sup>19</sup>, Jennifer A. Smit<sup>27</sup>, Marsden Solomon<sup>28,d</sup>, Lisa M. Soule<sup>29</sup>, Douglas Taylor<sup>2</sup>, Elizabeth E. Tolley<sup>2</sup>, Olivia Vandeputte<sup>30</sup>  
*\*Authors in alphabetical order from here onward*

### Affiliations

<sup>1</sup> Global Health and Population, FHI 360, Washington, District of Columbia, United States

<sup>2</sup> Global Health and Population, FHI 360, Durham, North Carolina, United States

<sup>3</sup> Department of Maternal and Child Health, University of North Carolina Gillings School of Global Public Health, Chapel Hill, North Carolina, United States

<sup>4</sup> Mann Global Health, Waterbury, Vermont, United States

<sup>5</sup> Johns Hopkins University School of Medicine and Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States

<sup>6</sup> Bill & Melinda Gates Foundation, Seattle, Washington, United States

<sup>7</sup> Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), National Institutes of Health (NIH), Bethesda, Maryland, United States

<sup>8</sup> Department of Obstetrics and Gynecology, University of Campinas, Faculty of Medical Sciences, Campinas, São Paulo, Brazil

<sup>9</sup> Profamilia, Santo Domingo, Dominican Republic

<sup>10</sup> University of Illinois Urbana-Champaign, Urbana, Illinois, United States

<sup>11</sup> Exeltis R&D, Madrid, Spain

<sup>12</sup> Research, Technology, and Utilization Division, Office of Population and Reproductive Health, Bureau for Global Health, United States Agency for International Development and the Public Health Institute, Washington, District of Columbia, United States

<sup>13</sup> University of California, Davis, Sacramento, California, United States

<sup>14</sup> Centre for Reproductive Health, Institute for Regeneration and Repair, University of Edinburgh, Edinburgh, Scotland, United Kingdom

<sup>15</sup> CONRAD - Eastern Virginia Medical School, Norfolk, Virginia, United States

<sup>16</sup> Oregon Health & Science University, Portland, Oregon, United States

<sup>17</sup> Bayer AG, Berlin, Germany

<sup>18</sup> Universität des Saarlandes, Homburg, Saarland, Germany

<sup>19</sup> Center for Biomedical Research, Population Council, New York, New York, United States

<sup>20</sup> Maternal, Child & Adolescent Health Program, Burnet Institute, Melbourne, Victoria, Australia

<sup>21</sup> Makerere University School of Public Health, Kampala, Uganda

<sup>22</sup> Women's Health Directorate, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, England, United Kingdom

<sup>23</sup> Global Health and Population, FHI 360, Atlanta, Georgia, United States

<sup>24</sup> University of Massachusetts Chan Medical School, Worcester, Massachusetts, United States

<sup>25</sup> College of Medicine, University of Ibadan, Ibadan, Oyo State, Nigeria

<sup>26</sup> Department of Obstetrics and Gynecology, Ribeirao Preto Medical School, University of São Paulo, Ribeirao Preto, São Paulo, Brazil

<sup>27</sup> Wits MatCH Research Unit, Department of Obstetrics and Gynaecology, The University of the Witwatersrand, Durban, KwaZulu-Natal, South Africa

<sup>28</sup> FHI 360, Nairobi, Kenya

<sup>29</sup> Independent Consultant, Bethesda, Maryland, United States

<sup>30</sup> MSI Reproductive Choices, London, England, United Kingdom

<sup>a</sup> Present address: Guttmacher Institute, New York, New York, United States

<sup>b</sup> Present address: Gedeon Richter Pharma, Berlin, Germany

<sup>c</sup> Present address: Independent Consultant, Washington, District of Columbia, United States

<sup>d</sup> Present address: Independent Consultant, Nairobi, Kenya

### **Corresponding author contact information**

Amelia Caroline Louise Mackenzie, PhD ScM  
2101 L St NW, Suite 700  
Washington, DC 20037 USA  
+1 919-321-3676  
amackenzie@fhi360.org

### **Acknowledgement of funding**

This expert consultation was made possible by the generous support of the American people through the U.S. Agency for International Development (USAID), provided to FHI 360 through cooperative agreement 7200AA20CA00016. The contents are the responsibility of FHI 360, the views expressed are those of the authors, and contents and views do not necessarily reflect those of any funder, including USAID, the National Institutes of Health, the United States Government, or the Reckitt Global Hygiene Institute Fellowship. ACLM, SC, EH, NM, AEB, RLC, AFC, ECC, LJD, AM, KN, MS, DT, and EET were supported but USAID cooperative agreement 7200AA20CA00016, with SC also supported by the Bill & Melinda Gates Foundation (OPP1200867) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (T32 HD52468) and MS also supported by the Bill & Melinda Gates Foundation (INV-045483). EC was supported by Exeltis, and CEN by USAID. LB, VB, KBHC, MDC, HODC, GFD, AE, LBH, JH, CRTJ, SPSK, DM, KAM, JAM, FMO, CBP, CSV, RSW, JAS, LMS, and OV received an honorarium supported by USAID cooperative agreement 7200AA20CA00016 that was declined by other authors. In addition, GFD was also supported by USAID Cooperative Agreement 7200AA20CA00019 provided to CONRAD/Eastern Virginia Medical School, JH was also supported by a Reckitt Global Hygiene Institute Fellowship, JAM was also supported by Wellcome Fellowship 209589/Z/17/Z, FMO was also supported by EVIHDAF, and RSW was also supported by The Population Council.

### **Conflicts of interest**

AEB has received research funding mediated through JHU from: Merck, Chemo Pharmaceuticals, Dare Bioscience, and NICHD. KBHC received a small honorarium from Pfizer to participate in their Tales From the Uterus podcast series in the summer/fall of 2023. EC is an Exeltis employee. MDC has received speaking honoraria from Gedeon Richter, Mayne, OLIC, and Organon, served on Advisory Boards for

Gedeon Richter and Mayne, has stock options with Femasys, and has consulted for Estetra SRL, Medicines360, and Organon. The Department of Obstetrics and Gynecology, University of California, Davis, receives contraceptive research funding for Dr. Creinin from Chemo Research SL, Evofem, Femasys, Medicines360, Merck, Sebela, and Sumitomo Pharma. HODC has received clinical research support for laboratory consumables and staff from Bayer AG (Paid to Institution) and has provided consultancy advice (Paid to Institution) for Bayer AG, PregLem SA, Gedeon Richter, Vifor Pharma UK Ltd, AbbVie Inc., Myovant Sciences GmbH, and Theramex. HODC has received royalties from UpToDate for an article on abnormal uterine bleeding. AE has received travel reimbursement from ACOG, WHO, CDC, and honoraria from Gynuity for committee activities. AE receives royalties from Up to Date, Inc. Oregon Health & Science University receives research funding from OHSU Foundation, Gates, Merck, HRA Pharma, and NIH where AE is the principal investigator. CG is an employee of Bayer AG. All other authors have no conflicts of interest to declare.

### **Word counts**

Abstract: 195 words

Text: 3,253 words

### **Running title**

Consensus recommendations for CIMCs in contraceptive clinical trials

### **Abstract**

Objective: We sought to develop consensus recommendations for improved measurement of contraceptive-induced menstrual changes (CIMCs) in contraceptive clinical trials to build upon previous standardization efforts.

Study Design: We undertook a virtual consensus-building process using a modified Delphi approach, incorporating aspects of Nominal Group Technique and the Jandhyala method. The process consisted of four half-day meetings, developing recommendations within five topical working groups, a series of consensus questionnaires to assess recommendation agreement on a 4-point Likert-like scale, and opportunities for input and feedback throughout the process. Recommendations required at least 75% consensus to be included.

Results: Our interdisciplinary group—from 26 organizations and 13 countries in five global regions across academia, nonprofit research organizations, the pharmaceutical industry, and funding agencies—developed 44 consensus recommendations. These included recommendations on standardization, patient-reported outcome measure development aligned with current regulatory guidance, trial design, protocol development, participant recruitment, data collection, data analysis, and areas for exploratory research. Seven recommendations were deemed a priority by over 85% of our group.

Conclusion: Using a formal consensus-building process, we reached high levels of agreement around recommendations for more standardized and patient-centered measurement of CIMCs in contraceptive clinical trials, now and in the future.

### **Keywords**

bleeding patterns; contraceptive-induced menstrual changes; contraceptive clinical trials; menstrual cycle; patient reported outcome measures; consensus

## Implications statement

Consensus recommendations on measuring bleeding changes and related outcomes in contraceptive clinical trials can improve reporting of standardized, patient-centered outcomes in future product labeling. These improvements can enable providers to offer more relevant counseling on contraceptives and permit greater comparability and data synthesis across trials to inform clinical guidance.

## 1. INTRODUCTION

The use of contraception can cause changes in uterine bleeding patterns, uterocervical fluid, and uterine cramping, and it can impact how users experience menstrual and gynecologic disorders and symptoms. Collectively, these changes can be referred to as contraceptive-induced menstrual changes (CIMCs), although there are differing perspectives on this terminology (Figure 1). Experiencing CIMCs can both negatively and positively impact use dynamics, health, and wellbeing for contraceptive users, as well as the individual and sociocultural acceptability of contraception [1]. When users find CIMCs undesirable, negative consequences can include contraceptive dissatisfaction, reduced quality of life, increased burden around managing menstrual health, reduced sexual wellness, and potential unintended pregnancy if users discontinue contraception while still wanting to prevent pregnancy [1–4]. On the other hand, experiences of CIMCs users consider desirable can result in benefits, such as management of menstrual and gynecologic disorders and symptoms, method satisfaction, improved sexual wellness, and reduced burden or costs of menstrual materials [1–6]. Which CIMCs users determine to be desirable or undesirable for them can vary widely and is influenced by individual preferences and norms and community-level norms—especially around menstruation, menstrual health, and sexual and reproductive health—as well as perspectives of partners and wider social and contextual factors [1–3,7,8]. However, broadly, users often prefer reduced bleeding and cramping, and dislike increased bleeding and unpredictability [1,2].

Figure 1. Terminology and definition of contraceptive-induced menstrual changes (CIMCs)

**The Global CIMC Task Force\* defines contraceptive-induced menstrual changes (CIMCs) as all the ways contraception can impact a user's former menstrual cycle, including:**

- Changes in bleeding duration, volume, frequency, and/or regularity/predictability;
- Changes in blood and other uterocervical fluid/effluent consistency, color, and/or smell;
- Changes in uterine cramping and pain;
- Changes in other symptoms formerly experienced during the menstrual cycle;
- Changes in experiences of menstrual and gynecologic disorders and symptoms, such as endometriosis, fibroids, and adenomyosis; and
- Short-term changes to the menstrual cycle after discontinuation.

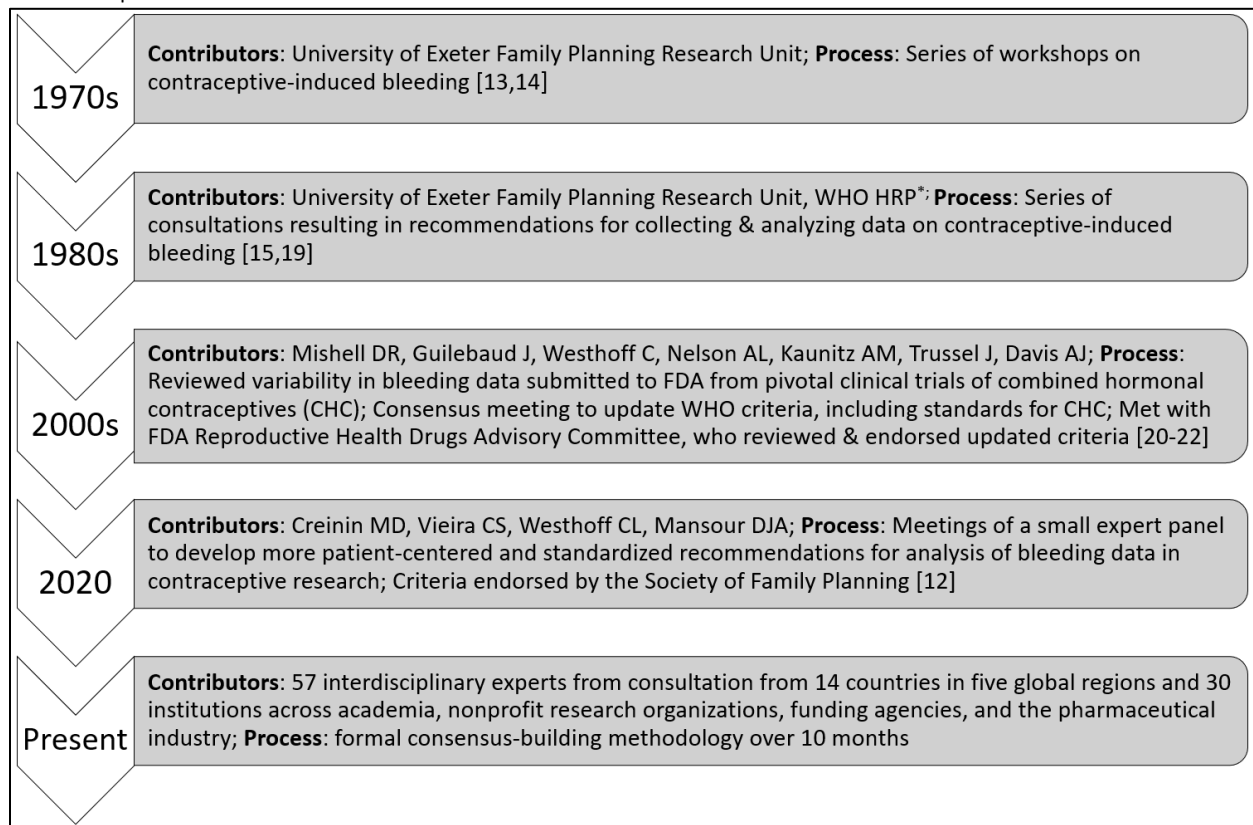
Terminology used for CIMCs differs depending on context. In this paper, we use the Global CIMC Task Forces' terminology and comprehensive definition because: (a) it captures all changes—not just changes in bleeding patterns—including changes in uterine cramping, the characteristics of uterocervical fluid, and other related symptoms; and (b) it is centered on how people who menstruate understand their own menstrual cycle when not using contraception and how that former cycle has been impacted by contraception.

Notably, however, the biological mechanisms underlying the endometrial changes that cause uterine bleeding while using hormonal contraception is not menstruation, and these users do not have a menstrual cycle. The framing of 'menstrual changes' in CIMCs in this case is intended to capture the contraceptive-induced changes *from* the menstrual cycles experienced before hormonal contraceptive use.

*\*Citation: Hoppes E, Nwachukwu C, Hennegan J, Blithe DL, Cordova-Gomez A, Critchley H, Doncel GF, Dorflinger LJ, Haddad LB, Mackenzie ACL, Maybin JA, Moley K, Nanda K, Vieira CS, Vwalika B, Kibira SPS, Mickler A, OlaOlorun FM, Polis CB, Sommer M, Williams KM, Lathrop E, Mahajan T, Rademacher KH, Solomon M, Wilson K, Wilson LC, Rountree L. Global research and learning agenda for building evidence on contraceptive-induced menstrual changes for research, product development, policies, and programs. *Gates Open Res.* 2022 Apr 19;6:49. doi: 10.12688/gatesopenres.13609.1.*

Considerations for CIMCs have played a notable role in contemporary contraceptive research and development (R&D). Beginning in the 1950s with the preclinical and clinical research for the first oral contraceptive, a number of development decisions were made, at least in part, due to CIMCs. For example, mestranol and a regimen of hormone-free days were both added because researchers decided monthly withdrawal bleeding was the most acceptable to contraceptive users, as well as possibly other contemporary influential stakeholders [9–11]. These decisions, made decades ago, still impact the current contraceptive method mix and contraceptive R&D today. Throughout this contemporary development of new contraceptives, there have been several endeavors to standardize how CIMCs are measured and reported, which we summarize in Figure 2. Creinin and colleagues have recently reviewed this history and the measurement outcomes from that work in detail [12]; therefore, we focus on the contributors and processes of these previous efforts to highlight differences from the approach reported in this paper.

Figure 2. Overview of efforts to standardize the measurement and analysis of CIMC data in contraceptive clinical trials



\* *Special Programme of Research, Development and Research Training in Human Reproduction* [23]

Current practices for measuring CIMCs in trials continue to have limitations, including burdensome data collection, a lack of uniform definitions provided to trial participants to understand key terms, and the use of patient-reported outcome (PRO) data without evidence of validation in line with current regulatory guidance on patient-focused drug development [24–27]. These limitations are particularly relevant for standard measurement of CIMCs in trials globally because—for topics like menstruation, with high levels of stigma—there are generally low levels of shared understanding and common references within and between communities and contexts [8,28,29]. In addition, how CIMC outcomes reported by trials for providers to use in contraceptive counseling of new methods are often not meaningful to contraceptive users to facilitate their informed method decision-making. Building upon previous standardization efforts—and in collaboration with many of the experts who contributed to this earlier work—we developed consensus recommendations to further improve how CIMCs are measured in contraceptive clinical trials now and in the future. Our consensus-building approach to address current measurement limitations focused on: (a) incorporating a variety of interdisciplinary and global perspectives, (b) considering current regulatory guidance around PROs, and (c) including all types of CIMCs, not just changes in bleeding patterns, to align with the Global CIMC Task Force’s comprehensive definition.

## 2. METHODS

Members of a scientific and technical team at FHI 360 engaged a group of interdisciplinary experts about participating in a consensus-building process to improve how CIMCs are measured in contraceptive clinical trials. Invitees had a range of expertise in topics (i.e., contraceptive clinical trials, family planning, the menstrual cycle, menstrual health, measurement and analysis approaches, and regulatory considerations) and disciplines (i.e., clinical research, social-behavioral science, biomedical research, basic science, and clinical practice). Experts were recruited within multiple sectors from members of the existing CIMC Global Task Force [1], contributors to previous standardization efforts, contributors to the literature in relevant fields, and suggestions from other experts. Among these experts, we convened a 13-member external planning committee from 12 organizations and 6 countries who volunteered to contribute to defining objectives for the expert consultation and refining the proposed consensus-building process.

Our objectives were: (a) to review and discuss limitations and strengths of current approaches for measuring and analyzing data on CIMCs in contraceptive clinical trials; (b) to come to a consensus on recommendations for improvements in CIMC measurement and analysis in contraceptive clinical trials that meet the needs of trial participants, researchers, sponsors, regulators, and future users; and (c) to identify important, related topics outside the scope of this expert consultation that warranted similar consideration. The second objective also included determining what research and evidence are needed to empirically evaluate the recommended improvements. Our consensus-building methodology was an amalgamation of the RAND/UCLA method, the Delphi method, the Modified Rand-Delphi method, the Nominal Group Technique, and the Jandhyala method because no single approach met all our needs [30–35]. We present extensive details on the methodology in the Supplementary Material, but briefly, the fully virtual consensus-building process, depicted in Figure 3, included four half-day meetings for all experts to join, recommendation development within five smaller topical working groups into which experts self-selected, and a series of consensus questionnaires around the scope and details of the recommendations.

Figure 3. Consensus-building methodology for developing recommendations on CIMC measurement in contraceptive clinical trials

STEP	PURPOSE
Day 1 meeting	Information sharing for common understanding*
Questionnaire 1	Define scope of recommendations (i.e., topics to be included)
Day 2 meeting	Consensus-building to finalize scope and define working groups
Working groups	Developing recommendations for each topic
Asynchronous learning	Filling knowledge gaps identified in Questionnaire 1 and meetings
Questionnaire 2	Feedback on recommendations drafted by Working Groups 1-3
Sync-Up Call	Discuss recommendation revisions and working group progress
Questionnaire 3	Feedback on recommendations from all working groups
Day 3 meeting	Working group presentations and consensus-building
Day 4 meeting	Working group presentations and consensus-building
Compile recommendations	Collate, streamline, and organize recommendations
Prioritization Survey	Final recommendation revision and prioritization

*\*To establish a shared understanding on key information within our interdisciplinary group, including on contraceptive clinical trials, patient-reported outcome measures, sociocultural contexts and CIMCs, and inclusive language around gender and sexuality*

Each consensus questionnaire asked experts about their agreement/disagreement with recommendations on a four-point Likert-like scale (i.e., strongly agree, somewhat agree, somewhat disagree, strongly disagree) and included at least one open-ended question for additional input. One author (SC) completed qualitative analysis of open-ended questions. During the first two days of meetings, we decided topics would need 85% agreement (i.e., ‘strongly agree’ or ‘somewhat agree’) among experts to be considered within scope of this work, and specific recommendations within those topics would need 75% agreement among members of working groups to be included in our final list. The Office of International Research Ethics at FHI 360 determined the expert consultation protocol was research that was exempt from Institutional Review Board review.

### 3. RESULTS

#### 3.1. Expert characteristics

A total of 57 experts were involved in the expert consultation from 30 organizations—across academia, nonprofit research organizations, funding agencies, and the pharmaceutical industry—and 14 countries in five global regions (i.e., Northern America, Europe, Africa, Latin America & the Caribbean, and Oceania, listed in order of frequency). Of those experts, 44 self-selected into working groups. Working group members represented over 625 combined years of experience—including nearly 70% with more than 15 years—in contraceptive clinical trials and other R&D, menstrual cycle and menstrual health research, social-behavioral research, contraceptive prescribing, and/or other relevant fields (Table 1). Most had doctoral training in either obstetrics and gynecology and/or research and worked either in academia or at nonprofit research organizations. Most experts self-identified as women, and about half

were under the age of 50 and half were 50 or more years old. Almost all experts (88%) reported having either experienced CIMCs themselves or observed someone personally close to them experiencing CIMCs.

Table 1: Characteristics of experts in topical working groups developing recommendations on CIMC measurement in contraceptive clinical trials

<b>Characteristic</b>	<b>Experts n (%)</b>
<b>Gender*</b>	
Woman	30 (73%)
Man	11 (27%)
<b>Age</b>	
Under 30	1 (2%)
30-39	7 (17%)
40-49	13 (32%)
50-59	8 (20%)
60-69	7 (17%)
70 or over	5 (12%)
<b>Area of Training</b>	
Research (PhD, ScD)	23 (56%)
Obstetrics and gynecology	14 (34%)
Practice (DrPh, MPH)	2 (5%)
Other	2 (5%)
<b>Primary Area(s) of Expertise**</b>	
Contraceptive clinical trials/R&D	24 (59%)
Social-behavioral research	17 (41%)
Menstrual cycle/menstrual health research	15 (37%)
Other contraceptive research	16 (39%)
Contraceptive provider/prescriber	12 (29%)
Analysis methodologies	7 (17%)
Patient-reported outcome measures	5 (12%)
Regulatory approval	4 (10%)
<b>Years of experience</b>	
6-10 years	7 (17%)
11-15 years	6 (15%)
16-20 years	11 (27%)
More than 20 years	17 (41%)
<b>Primary affiliation</b>	
Academic institution	15 (37%)
Non-governmental organization (NGO)	14 (34%)
Funder agency/organization	6 (15%)
Consultant	3 (7%)
Pharmaceutical industry	2 (5%)
<b>Total</b>	<b>41†</b>

\*Self-identified gender as woman, man, nonbinary or gender diverse, or self-described.



\*\* Experts could select up to three areas, so the sum is greater than 41 (100%).

† Three working groups members did not provide data for Table 1

### 3.2. Defining the scope for recommendation development

All 11 topics proposed in Questionnaire 1 reached consensus for inclusion (Table 2). Experts suggested another five topics that did not reach consensus during voting during the Day 2 meeting, but a number of elements of those topics were incorporated into other consensus topics. Topics determined to be out of scope are presented in Supplementary Table S1, aligning with the third objective of the expert consultation to identify important, related topics that warrant similar consideration outside of the present process. We converged the 11 topics into 3 cross-cutting themes and 5 topical working groups (Table 2). Working group scopes were: (1) **Eligibility**, enrollment requirements, and confounders; (2) Type, frequency, and format/mode of CIMC **data** collected; (3) Measurement of **acceptability**, impact on daily life, culture, and context; (4) Research agenda for developing CIMC **instruments**; (5) **Analysis** methodology, measures, and analyses standardization.

Table 2. Scope of topics for recommendations on CIMC measurement in contraceptive clinical trials, as determined via Questionnaire 1 and Day 2 meeting, and the corresponding topical working groups

Topics from Questionnaire 1	Percent agreement	Related working group
Eligibility and enrollment requirements	92%	Working Group 1. Eligibility
Data collection on confounders and other variables	95%	
Frequency of CIMC data collection	100%	Working Group 2. Data
CIMC instrument format and mode	95%	
Type of CIMC data collected during clinical trials	100%	
Measurement of acceptability and impact on daily life	100%	Working Group 3. Acceptability
Culture and context	87%	
Developing new CIMC instruments	85%	Working Group 4. Instruments
Measures and analyses standardization	97%	Working Group 5. Analysis
Analysis methodology	85%	
Regulatory considerations	85%	Cross-Cutting Themes
Culture and context	87%	
Improved counseling information *†	56%	
Guiding frameworks (reproductive justice, life course, equity) *	47%	Did not reach consensus during Day 2 voting.
Trial sample characterization (life stage, fertility intentions, goals/expectations, desires/needs) *	50%	
Non-contraceptive use (label information, menstrual choice, impact on wider health) *	47%	
Menstrual changes data in other trials beyond contraception *	47%	

\* Suggested by experts in Questionnaire 1, with percent agreement from Day 2 voting

† Revised into “Incorporating User Perspectives and Understandings of CIMCs” and agreed upon as cross-cutting theme

### 3.3. Final Consensus Recommendations

We organized the final 44 consensus recommendations into six sections on: (a) *Standardization* in recommendations 1.1 to 1.4; (b) *Instruments for measuring CIMCs* in 2.1 to 2.3; (c) *Trial design, protocol development, and participant recruitment* in 3.1 to 3.8; (d) *Data collection* in 4.1 to 4.13; (e) *Data analysis* in 5.1 to 5.14; and (f) *Areas for exploratory research* in 6.1 and 6.2 (Table 3). We reached

consensus on all recommendations, with nearly half (n=21) achieving greater than 90% consensus and six reaching 100% consensus (see Supplementary Table S2 for full consensus results). One aspect of one recommendation did not reach the consensus threshold of 75% by the end of our process (i.e., part of 5.3 on variables for analysis, details in Table S2), so it was not included. During the final stage of the consensus-building process (Figure 3), we identified priority recommendations based on the urgency and potential impact of their implementation (see Supplementary Table S2 for additional details on consensus and prioritization outcomes). This included seven recommendations that over 85% of experts deemed to be a priority and an additional 12 that at least 75% deemed a priority (highlighted by “\*” or “+”, respectively in Table 3).

Our final recommendations include research approaches to establish the evidence for future improvement in CIMC measurement in trials (e.g., 2.1, 2.2 and 2.3) and recommendations investigators can implement in the meantime, while that research is being conducted (i.e., recommendations beginning with “presently...”, e.g., 4.1 through 4.13 and 5.2 through 5.12). As our aim was to build upon, and not replicate, previous work standardizing CIMC measurement in clinical trials, recommendations often incorporate concepts from previous criteria, especially the work of Creinin and colleagues [12]. Recommendations 1.3 and 5.2 draw heavily from these most recent criteria. Several recommendations also include and refer to regulatory guidance and best practices around developing measures for PROs like CIMCs.

Table 3. Consensus recommendations on CIMC measurement in contraceptive clinical trials  
[See Table 3 at end of document.](#)

## 4. DISCUSSION

Building upon previous approaches at standardization around CIMCs, our consensus recommendations were developed by virtually convening interdisciplinary experts from around the world to consider how to improve the current and future measurement and analysis of CIMC data in contraceptive clinical trials. The resulting 44 recommendations, including identified priorities, can serve as guidance to those implementing and funding clinical trials. Based on the thoughtful input throughout the expert consultation, there are a few important considerations we wish to highlight for those implementing our recommendations. We also note reflections in the Supplementary Material on our consensus-building methodology, which we propose as a model—for the field of contraception, wider sexual and reproductive health, and beyond—of how to successfully bring a large, interdisciplinary, and diverse group of people together from around the world to come to consensus on a topic that had eluded such previous agreement.

### 4.1 Reflections on developing recommendations

Although we ultimately reached consensus on almost every aspects of our recommendations, a few areas of contention emerged during the process. In some instances, differences occurred simply due to the interdisciplinary and global nature of our expert group and the distinct perspectives that inherently arise when convening those with different trainings, experiences, and areas of expertise. Often these differences emerged between those approaching the work from a biomedical perspective verses those working within social-behavioral research. All these differences and the ensuing conversations, however, were essential to thoroughly addressing a topic as multifaceted as measuring CIMCs in the context of clinical trials.

Some areas of contention proved especially challenging and often could be distilled down to differing perspectives along a continuum. On one end of the spectrum was a preference to recommend modest improvements to current approaches that would be feasible and straightforward to implement, aligning with past and present norms and existing systems; on the other end was a desire to push for larger, creative improvements aimed at innovating beyond current confines. In addition, we also encountered an innate tension in attempting to weigh the often-divergent perspectives of various stakeholders within the contraceptive clinical trial ecosystem, including trial investigators and other researchers, trial sponsors and funders, regulators, trial participants, and the future providers and users of the new contraceptives being developed in trials. Because we were developing recommendations for global use, balancing these different perspectives was even more challenging considering the wide range of contexts in which clinical trials are conducted, including within different healthcare systems, countries, and resource settings around the world. We describe a few examples of weighing these differing perspectives next.

Recommendations needed to be specific and technical enough to provide actionable guidance to trial investigators without being too burdensome to enact within the many constraints and competing demands of conducting a clinical trial. From this researcher perspective, there can be an appeal of smaller, incremental improvements to current practices that could be simple to implement with often-limited resources. This gradual approach could also be appealing from a sponsor and funder standpoint because it may not require increased trial budgets and may be more likely aligned with current and familiar regulatory strategies. For some experts, these types of modest changes within current norms and systems were of interest for very practical, realistic reasons founded in their many years of experience. Indeed, complex and/or numerous novel recommendations that may demand additional time and funding could result in delaying upcoming trials, as well as fewer, longer trials in the future. In addition, many experts were concerned our recommendations would be interpreted by funders or regulators as expert-endorsed mandates that should be required for every trial. Altogether, the broader impact of these types of additional burdens and constraints could be fewer new methods being developed, which could ultimately mean the contraceptive needs of some people and couples may not be met. To address the concerns our recommendations could hinder contraceptive R&D, we specify they should not be considered required for contraceptive trials to be funded, conducted, or reviewed by regulatory authorities. Rather, stakeholders should consider these as recommended approaches to achieving standardization and comparability across trials and to guiding future research for improvement, all with emphasis on identified priorities.

On the other hand, this convening of around 50 interdisciplinary experts from multiple global regions in a consensus-building process was a unique opportunity to push the contraceptive clinical trial ecosystem to consider innovative CIMC data collection and analysis approaches that could be more informative to providers in counseling on new contraceptives and more relevant to future users. Many experts endorsed this perspective, which is also aligned with current patient-focused drug development initiatives, other efforts at increasing patient engagement in clinical trials and regulatory decisions, and wider patient-centered outcome research [24–27,36,37]. As described in our recommendations, research will be needed to align CIMC measurement with best practices for PROs. Because this approach is novel for contraceptive trials—although increasingly common in other therapeutic areas—it will require close adherence to regulatory guidelines and engagement with regulators.

Although areas of contention persisted, our approach sought to balance the perspectives of all experts in the consensus-building process and stakeholders in the contraceptive clinical trial ecosystem in our final recommendations by focusing on the goal of improving the standardization and value of trial data on CIMCs. For example, this is the first time recommendations have included uterine cramping and pain, an important outcome for people who menstruate. By considering all CIMCs, not just changes in bleeding patterns, and including experts in menstrual health, heavy menstrual bleeding, patient-reported outcomes, social-behavioral researchers, and others, we were able to extend the boundaries of this work. One related limitation to our process we must acknowledge is the lack of direct input from clinical trial participants. Although we aimed to include their perspectives, as understood among our expert group, this does not replace explicit contributions from participants and their communities. To ensure the utility and relevance of the recommendations around future research a balanced and inclusive approach will need to continue, including using participatory-engaged research methodologies.

#### **4.2 Implementing recommendations**

As a starting point for implementing these recommendations, we have highlighted seven priorities: 1.1, 2.2, 4.1, 4.11, 5.2, 5.3, and 5.11 (listed separately in Supplementary Table S3). Recommendations 1.1, 2.2, and 5.2 also reached 100% consensus. Recommendation 1.1 calls for standardization and simplification of terminology. In support of this, 2.2 outlines best practices for making both terminology and measurement instruments simple, accessible, and patient-centered. Recommendation 5.2, also aligning with 1.1, lays out important terminology and definitions for use across trials, defining a bleeding episode, as well as duration, frequency, and volume of bleeding episodes. Recommendations 1.1 also calls for standardization in the type of CIMC data collected; therefore, 4.1 and 4.11 provide more detail on data collection, with 4.1 covering the primary outcomes of CIMCs that should be collected, including bleeding and spotting days and frequency, duration, and volume of bleeding episodes, and 4.11 outlining the acceptability and quality of life data to be collected. Finally, 5.3 and 5.11 outline some key variables and missing data that we recommend be reported for all contraceptive clinical trials.

Even implementing only these seven priority recommendations will come with challenges. We recognize the list of all recommendations in Table 3 is not particularly accessible or user-friendly to all. An important next step is to develop tools for stakeholders across the contraceptive R&D ecosystem, such as checklists and guidelines to make adopting these recommendations more straight-forward without imposing any unnecessary burden on the development of new contraceptives. It will also be valuable to tailor these resources to the types of research being conducted (e.g., Phase 1 vs. Phase 3, or clinical vs. social-behavioral research) and to target audiences (e.g., researchers, funders, providers, regulators).

#### **4.3 Future implications**

We call on stakeholders in the contraceptive clinical trial ecosystem to view these recommendations as new best practices, developed and endorsed by a wide range of experts in contraceptive R&D and related disciplines. We encourage those involved in the design of clinical trials to begin to incorporate these recommendations into their work as soon as possible. To achieve the full benefits of these recommendations, additional dissemination will be required. This publication is only the beginning of efforts needed to make these recommendations actual best practices amongst the contraceptive R&D community.

We also note contraceptives are not the only drug, device, or biologic to impact the menstrual cycle, yet such data are not routinely collected during clinical trials or during standard toxicology and

pharmacodynamics studies, as other organ functions and vital signs are. The lack of these data was recently highlighted with the initial introduction of COVID vaccinations when vaccinated people who could or did menstruate experienced unanticipated bleeding changes [38–42]. We hope our work can be a guide and an invitation to develop recommendations for if, when, and how to measure changes to the menstrual cycle in more types of clinical trials and related research.

Broadly, our consensus recommendations can enable more standardized and patient-centered measurement and analysis of CIMC data in the near-term and continued improvement in the future, which can result in more accurate information in contraceptive product labeling on data that matter more to users. Ultimately, such efforts can allow for greater comparability across trials and better data synthesis in systematic reviews and meta-analyses that can inform clinical guidance. In turn, these improvements can enable providers to offer better counseling and users to make more informed decisions about their choice of method, which can improve the experiences of contraceptive users and better meet the contraceptive needs of people and couples.

### Supplementary material

1. Supplementary Material file
2. Supplementary tables

### Acknowledgments

This expert consultation was made possible by the generous support of the American people through the U.S. Agency for International Development (USAID), provided to FHI 360 through cooperative agreement 7200AA20CA00016. The contents are the responsibility of FHI 360, the views expressed are those of the authors, and contents and views do not necessarily reflect those of any funder, including USAID, the National Institutes of Health, the United States Government, or the Reckitt Global Hygiene Institute Fellowship. ACLM, SC, EH, NM, AEB, RLC, AFC, ECC, LJD, AM, KN, MS, DT, and EET were supported but USAID cooperative agreement 7200AA20CA00016, with SC also supported by the Bill & Melinda Gates Foundation (OPP1200867) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (T32 HD52468) and MS also supported by the Bill & Melinda Gates Foundation (INV-045483). EC was supported by Exeltis, and CEN by USAID. LB, VB, KBHC, MDC, HODC, GFD, AE, LBH, JH, CRTJ, SPSK, DM, KAM, JAM, FMO, CBP, CSV, RSW, JAS, LMS, and OV received an honorarium supported by USAID cooperative agreement 7200AA20CA00016 that was declined by other authors. In addition, GFD was also supported by USAID Cooperative Agreement 7200AA20CA00019 provided to CONRAD/Eastern Virginia Medical School, JH was also supported by a Reckitt Global Hygiene Institute Fellowship, JAM was also supported by Wellcome Fellowship 209589/Z/17/Z, FMO was also supported by EVIHDAF, and RSW was also supported by The Population Council.

The FHI 360 team would like to recognize the additional contributions of the external planning committee members, whose feedback greatly improved the consensus-building process. These members are Amanda Cordova-Gomez, Anita Nelson, Bellington Vwalika, Carolina Sales Vieira, Chelsea Polis, Diana Mansour, Funmi OlaOlorun, Kathryn Clancy, Kristen Matteson, Lisa Haddad, Mitchell Creinin, Olivia Vandeputte, and Tanya Dargan Mahajan. We also appreciate the *ad hoc* advisory roles of Lisa Rarick on regulatory considerations, Marni Sommer on menstrual health, and Greg Kopf on preclinical and early pharmaceutical trials. We thank Leigh Wynne for input on virtual meeting engagement approaches and coordinating Working Group 5. We also appreciate feedback from Aurélie Brunie and Holly Burke on the

manuscript. Natt Pladsri and Audrey Fratus provided important support in preparing for the virtual meetings and ensuring they ran smoothly to aid in successful discussion.

## References

- [1] Hoppes E, Nwachukwu C, Hennegan J, Blithe DL, Cordova-Gomez A, Critchley H, et al. Global research and learning agenda for building evidence on contraceptive-induced menstrual changes for research, product development, policies, and programs. *Gates Open Res* 2022;6:49. <https://doi.org/10.12688/gatesopenres.13609.1>.
- [2] Polis CB, Hussain R, Berry A. There might be blood: A scoping review on women’s responses to contraceptive-induced menstrual bleeding changes. *Reprod Health* 2018;15:1–17. <https://doi.org/10.1186/s12978-018-0561-0>.
- [3] Higgins JA, Smith NK. The Sexual Acceptability of Contraception: Reviewing the Literature and Building a New Concept. *J Sex Res* 2016;53:417–56. <https://doi.org/10.1080/00224499.2015.1134425>.
- [4] Higgins JA, Davis AR. Contraceptive sex acceptability: a commentary, synopsis and agenda for future research. *Contraception* 2014;90:4–10. <https://doi.org/10.1016/j.contraception.2014.02.029>.
- [5] Brunie A, Stankevitz K, Nwala AA, Nqumayo M, Chen M, Danna K, et al. Expanding long-acting contraceptive options: a prospective cohort study of the hormonal intrauterine device, copper intrauterine device, and implants in Nigeria and Zambia. *Lancet Glob Health* 2021;9:e1431–41. [https://doi.org/10.1016/S2214-109X\(21\)00318-1](https://doi.org/10.1016/S2214-109X(21)00318-1).
- [6] Rademacher KH, Sripipatana T, Danna K, Sitrin D, Brunie A, Williams KM, et al. What Have We Learned? Implementation of a Shared Learning Agenda and Access Strategy for the Hormonal Intrauterine Device. *Glob Health Sci Pract* 2022;10. <https://doi.org/10.9745/GHSP-D-21-00789>.
- [7] Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017;17:88. <https://doi.org/10.1186/s12913-017-2031-8>.
- [8] Tan DA, Haththotuwa R, Fraser IS. Cultural aspects and mythologies around menstruation and abnormal uterine bleeding. *Best Pract Res Clin Obstet Gynaecol* 2016:1–13. <https://doi.org/10.1016/j.bpobgyn.2016.09.015>.
- [9] Speroff L, Darney PD. The history of contraception. In: Jensen JT, Creinin MD, editors. *Speroff & Darney’s Clinical Guide to Contraception*. 6th ed., Philadelphia, Pennsylvania: Wolters Kluwer; 2020, p. online chapter.
- [10] Drucker DJ. *Contraception: a concise history*. Cambridge, Massachusetts: Massachusetts Institute of Technology Press; 2020.
- [11] Junod SW, Marks L. Women’s trials: the approval of the first oral contraceptive pill in the United States and Great Britain. *Journal of the History of Medicine* 2002;57:117–60. <https://doi.org/10.1093/jhmas/57.2.117>.

- [12] Creinin MD, Vieira CS, Westhoff CL, Mansour DJA. Recommendations for standardization of bleeding data analyses in contraceptive studies. *Contraception* 2022;112:14–22. <https://doi.org/10.1016/j.contraception.2022.05.011>.
- [13] Snowden R. The statistical analysis of menstrual bleeding patterns. *J Biosoc Sci* 1977;9:107–20. <https://doi.org/10.1017/s0021932000000511>.
- [14] Rodriguez G, Faundes-Latham A, Atkinson LE. An approach to the analysis of menstrual patterns in the critical evaluation of contraceptives. *Stud Fam Plann* 1976;7:42–51.
- [15] Belsey EM, Machin D, D’Arcangues C. The analysis of vaginal bleeding patterns induced by fertility regulating methods. World Health Organization Special Programme of Research, Development and Research Training in Human Reproduction. *Contraception* 1986;34:253–60. [https://doi.org/10.1016/0010-7824\(86\)90006-5](https://doi.org/10.1016/0010-7824(86)90006-5).
- [16] Belsey EM, Farley TM. The analysis of menstrual bleeding patterns: a review. *Contraception* 1988;38:129–56.
- [17] Machin D, D’arcangues C, Busca B, Farley TMM, Pinol A. Vaginal bleeding patterns — the problem and an example data set. *Applied Stochastic Models and Data Analysis* 1987;3:27–35. <https://doi.org/10.1002/asm.3150030104>.
- [18] Belsey EM, Farley TMM. The analysis of menstrual bleeding patterns: A review. *Applied Stochastic Models and Data Analysis* 1987;3:125–50. <https://doi.org/10.1002/asm.3150030302>.
- [19] Belsey EM, Carlson N. The description of menstrual bleeding patterns: Towards fewer measures. *Stat Med* 1991;10:267–84. <https://doi.org/10.1002/sim.4780100210>.
- [20] Mishell DR, Guillebaud J, Westhoff C, Nelson AL, Kaunitz AM, Trussell J, et al. Combined hormonal contraceptive trials: variable data collection and bleeding assessment methodologies influence study outcomes and physician perception. *Contraception* 2007;75:4–10. <https://doi.org/10.1016/j.contraception.2006.08.008>.
- [21] Mishell DR, Guillebaud J, Westhoff C, Nelson AL, Kaunitz AM, Trussell J, et al. Recommendations for standardization of data collection and analysis of bleeding in combined hormone contraceptive trials. *Contraception* 2007;75:11–5. <https://doi.org/10.1016/j.contraception.2006.08.012>.
- [22] US Food and Drug Administration. Advisory Committee for Reproductive Health Drugs final summary meeting minutes 2007. <https://wayback.archive-it.org/7993/20170404050830/https://www.fda.gov/ohrms/dockets/ac/07/minutes/2007-4274m1.pdf> (accessed December 6, 2023).
- [23] World Health Organization. HRP at 50: The UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP) 2023. [https://www.who.int/teams/sexual-and-reproductive-health-and-research-\(srh\)/human-reproduction-programme/hrp-at-50](https://www.who.int/teams/sexual-and-reproductive-health-and-research-(srh)/human-reproduction-programme/hrp-at-50) (accessed December 6, 2023).
- [24] US Food and Drug Administration. Patient-Focused Drug Development: Collecting Comprehensive and Representative Input: Guidance for Industry, Food and Drug Administration Staff, and Other

- Stakeholders 2020. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-collecting-comprehensive-and-representative-input> (accessed December 6, 2023).
- [25] US Food and Drug Administration. Patient-Focused Drug Development: Methods to Identify What Is Important to Patients: Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders 2022. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-methods-identify-what-important-patients> (accessed December 6, 2023).
- [26] US Food and Drug Administration. Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments: Draft Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders 2022. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-selecting-developing-or-modifying-fit-purpose-clinical-outcome> (accessed December 6, 2023).
- [27] US Food and Drug Administration. Patient-Focused Drug Development: Incorporating Clinical Outcome Assessments Into Endpoints for Regulatory Decision-Making 2023. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-incorporating-clinical-outcome-assessments-endpoints-regulatory> (accessed December 6, 2023).
- [28] Johnston-Robledo I, Chrisler JC. The Menstrual Mark: Menstruation as Social Stigma. In: Bobel C, Winkler I, Fahs B, Hasson K, Kissling E, Roberts T, editors. *The Palgrave Handbook of Critical Menstruation Studies*, Singapore: Springer Singapore; 2020, p. 181–99. [https://doi.org/10.1007/978-981-15-0614-7\\_17](https://doi.org/10.1007/978-981-15-0614-7_17).
- [29] Das M. Socio-cultural aspects of menstruation: An anthropological purview. *East Anthropol* 2008;61:227–40.
- [30] Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. *The RAND/UCLA Appropriateness Method User’s Manual*. Santa Monica, CA: 2001.
- [31] Broder MS, Gibbs SN, Yermilov I. An Adaptation of the RAND/UCLA Modified Delphi Panel Method in the Time of COVID-19. *J Healthc Leadersh* 2022;14:63–70. <https://doi.org/10.2147/JHL.S352500>.
- [32] Beretta R. A critical review of the Delphi technique. *Nurse Res* 1996;3:79–89. <https://doi.org/10.7748/nr.3.4.79.s8>.
- [33] Jandhyala R. Delphi, non-RAND modified Delphi, RAND/UCLA appropriateness method and a novel group awareness and consensus methodology for consensus measurement: a systematic literature review. *Curr Med Res Opin* 2020;36:1873–87. <https://doi.org/10.1080/03007995.2020.1816946>.
- [34] Horton JN. Nominal group technique. A method of decision-making by committee. *Anaesthesia* 1980;35:811–4. <https://doi.org/10.1111/j.1365-2044.1980.tb03924.x>.



- [35] Jandhyala R. A novel method for observing proportional group awareness and consensus of items arising from list-generating questioning. *Curr Med Res Opin* 2020;36:883–93. <https://doi.org/10.1080/03007995.2020.1734920>.
- [36] Gnanasakthy A, Qin S, Norcross L. FDA Guidance on Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments: Old Wine in a New Bottle? *The Patient - Patient-Centered Outcomes Research* 2023;16:3–5. <https://doi.org/10.1007/s40271-022-00607-6>.
- [37] Patient involvement in the development, regulation and safe use of medicines. Geneva: 2022. <https://doi.org/10.56759/iiew8982>.
- [38] Lee KMN, Junkins EJ, Luo C, Fatima UA, Cox ML, Clancy KBH. Investigating trends in those who experience menstrual bleeding changes after SARS-CoV-2 vaccination. *Sci Adv* 2022;8:eabm7201. <https://doi.org/10.1126/sciadv.abm7201>.
- [39] Edelman A, Boniface ER, Benhar E, Han L, Matteson KA, Favaro C, et al. Association Between Menstrual Cycle Length and Coronavirus Disease 2019 (COVID-19) Vaccination: A U.S. Cohort. *Obstetrics and Gynecology* 2022. <https://doi.org/10.1097/AOG.0000000000004695>.
- [40] Gibson EA, Li H, Fruh V, Gabra M, Asokan G, Jukic AMZ, et al. Covid-19 vaccination and menstrual cycle length in the Apple Women’s Health Study. *NPJ Digit Med* 2022;5:165. <https://doi.org/10.1038/s41746-022-00711-9>.
- [41] Wesselink AK, Lovett SM, Weinberg J, Geller RJ, Wang TR, Regan AK, et al. COVID-19 vaccination and menstrual cycle characteristics: A prospective cohort study. *Vaccine* 2023;41:4327–34. <https://doi.org/10.1016/j.vaccine.2023.06.012>.
- [42] Kareem R, Sethi MR, Inayat S, Irfan M. The effect of COVID-19 vaccination on the menstrual pattern and mental health of the medical students: A mixed-methods study from a low and middle-income country. *PLoS One* 2022;17:e0277288. <https://doi.org/10.1371/journal.pone.0277288>.
- [43] Mackenzie A, Chung S, Hoppes E, Mickler A, Cartwright A. Measurement of changes to the menstrual cycle: A transdisciplinary systematic review evaluating measure quality and utility for clinical trials. Submitted to *PLoS One* 2023.
- [44] USG General Services Administration. Federal plain language guidelines. 2023 n.d. <https://www.plainlanguage.gov/guidelines/> (accessed December 6, 2023).
- [45] World Wide Web Consortium (W3C). Web Content Accessibility Guidelines 2023. <https://www.w3.org/standards/> (accessed December 6, 2023).
- [46] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8:94–104. <https://doi.org/10.1111/j.1524-4733.2005.04054.x>.
- [47] Wild D, Eremenco S, Mear I, Martin M, Houchin C, Gawlicki M, et al. Multinational trials- recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported

- Outcomes Translation and Linguistic Validation Good Research Practices Task Force report. *Value Health* 2009;12:430–40. <https://doi.org/10.1111/j.1524-4733.2008.00471.x>.
- [48] Eremenco S, Coons SJ, Paty J, Coyne K, Bennett A V, McEntegart D, et al. PRO data collection in clinical trials using mixed modes: report of the ISPOR PRO mixed modes good research practices task force. *Value Health* 2014;17:501–16. <https://doi.org/10.1016/j.jval.2014.06.005>.
- [49] Byrom B, Gwaltney C, Slagle A, Gnanasakthy A, Muehlhausen W. Measurement Equivalence of Patient-Reported Outcome Measures Migrated to Electronic Formats: A Review of Evidence and Recommendations for Clinical Trials and Bring Your Own Device. *Ther Innov Regul Sci* 2019;53:426–30. <https://doi.org/10.1177/2168479018793369>.
- [50] Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: Addressing ethical challenges. *PLoS Med* 2018;15:e1002689. <https://doi.org/10.1371/journal.pmed.1002689>.
- [51] Mooney SJ, Pejaver V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. *Annu Rev Public Health* 2018;39:95–112. <https://doi.org/10.1146/annurev-publhealth-040617-014208>.
- [52] Price WN, Cohen IG. Privacy in the age of medical big data. *Nat Med* 2019;25:37–43. <https://doi.org/10.1038/s41591-018-0272-7>.
- [53] Jeyaraman M, Balaji S, Jeyaraman N, Yadav S. Unraveling the Ethical Enigma: Artificial Intelligence in Healthcare. *Cureus* 2023;15:e43262. <https://doi.org/10.7759/cureus.43262>.
- [54] Abramoff MD, Tarver ME, Loyo-Berrios N, Trujillo S, Char D, Obermeyer Z, et al. Considerations for addressing bias in artificial intelligence for health equity. *NPJ Digit Med* 2023;6:170. <https://doi.org/10.1038/s41746-023-00913-9>.
- [55] 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 Gynaecol Obstet* 2018;143:393–408. <https://doi.org/10.1002/ijgo.12666>.
- [56] de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge University Press; 2011. <https://doi.org/10.1017/CBO9780511996214>.
- [57] Severy LJ, Newcomer S. Critical issues in contraceptive and STI acceptability research. *Journal of Social Issues* 2005;61:45–65. <https://doi.org/10.1111/j.0022-4537.2005.00393.x>.
- [58] Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905. <https://doi.org/10.1007/s11136-012-0344-y>.
- [59] Gutin I. In BMI We Trust: Reframing the Body Mass Index as a Measure of Health. *Soc Theory Health* 2018;16:256–71. <https://doi.org/10.1057/s41285-017-0055-0>.
- [60] Bray GA. Beyond BMI. *Nutrients* 2023;15. <https://doi.org/10.3390/nu15102254>.

- [61] Gutin I. Body mass index is just a number: Conflating riskiness and unhealthiness in discourse on body size. *Sociol Health Illn* 2021;43:1437–53. <https://doi.org/10.1111/1467-9566.13309>.
- [62] Deurenberg P, Deurenberg-Yap M. Validity of body composition methods across ethnic population groups. *Forum Nutr* 2003;56:299–301.
- [63] US Food and Drug Administration. Establishing Effectiveness and Safety for Hormonal Drug Products Intended to Prevent Pregnancy Guidance for Industry 2019. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/establishing-effectiveness-and-safety-hormonal-drug-products-intended-prevent-pregnancy-guidance> (accessed December 6, 2023).
- [64] Lin G, Rodriguez R. Weighted Methods for Analyzing Missing Data with the GEE Procedure. *SAS Conference Proceedings: SAS Global Forum 2014*, 2014.
- [65] SAS Institute Inc. PROC GEE. *SAS/STAT® 15.2 User's Guide*, Cary, North Carolina: SAS Institute Inc.; 2020.
- [66] Fitzmaurice GM, Molenberghs G, Lipsitz SR. Regression Models for Longitudinal Binary Responses with Informative Drop-Outs. *Journal of the Royal Statistical Society: Series B (Methodological)* 1995;57:691–704. <https://doi.org/10.1111/j.2517-6161.1995.tb02056.x>.
- [67] Henriksen-Bulmer J, Jeary S. Re-identification attacks—A systematic literature review. *Int J Inf Manage* 2016;36:1184–92. <https://doi.org/10.1016/j.ijinfomgt.2016.08.002>.
- [68] Chikwetu L, Miao Y, Woldetensae MK, Bell D, Goldenholz DM, Dunn J. Does deidentification of data from wearable devices give us a false sense of security? A systematic review. *Lancet Digit Health* 2023;5:e239–47. [https://doi.org/10.1016/S2589-7500\(22\)00234-5](https://doi.org/10.1016/S2589-7500(22)00234-5).
- [69] Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y. Identifying personal genomes by surname inference. *Science* 2013;339:321–4. <https://doi.org/10.1126/science.1229566>.
- [70] Tyndall A, Du W, Breder CD. Regulatory watch: The target product profile as a tool for regulatory communication: advantageous but underused. *Nat Rev Drug Discov* 2017;16:156. <https://doi.org/10.1038/nrd.2016.264>.

1 **Table 3. Consensus<sup>i</sup> recommendations on CIMC measurement in contraceptive clinical trials**

**Table 3 Key**

**	>85% of experts agree that this recommendation should be prioritized (See Supplementary Table S3)
*	75%-84% of experts agree that this recommendation should be prioritized
†	Recommendation adapted from Creinin et al., 2022 [12]
§	Recommendation informed by regulatory guidance or best practices on measuring patient-reported outcomes (PROs), like CIMCs

Number	Recommendation
<b>1. Standardization recommendations</b>	
1.1 ** §	As much as possible, research on CIMCs should aim to: <ol style="list-style-type: none"> <li>a. Simplify terminology and/or align terminology with established standards to ease and improve global communication and collaboration around CIMCs in clinical trials among various stakeholders.<sup>ii</sup></li> <li>b. Collect data and describe CIMCs in a standardized way to allow comparisons across studies and products; researchers may also include assessments of participants’ perceptions of and experience with CIMCs.</li> <li>c. Use standardized simple, patient-centered terminology when reporting CIMC data. This terminology should be based on evidence and not be overly prescriptive. In addition, researchers and analysts should include previous analysis approaches and terminology when necessary to permit comparisons to data from previous studies of existing methods.</li> </ol>
1.2	Researchers should ideally consider how any new instruments <sup>iii</sup> used to collect CIMC data relate to commonly used questions in previous trials of already approved contraceptive methods—even if those questions were not part of a validated instrument—to allow comparability of data between current contraceptive methods and new methods for providers, users, and researchers, including for future systematic reviews. <i>Also See 2.1 to 2.3.</i>

<sup>i</sup> The consensus threshold was 75% agreement (i.e., responding “strongly agree” or “somewhat agree”), aligned with recommended Delphi methodology [30]. The level of consensus reached for each recommendation and agreement on priorities are reported in Supplementary Table S2.

<sup>ii</sup> These standards can include FIGO’s System 1 for nomenclature of normal and abnormal uterine bleeding [55], the Global CIMC Task Force’s definition of CIMCs [1], recommendations on bleeding data analyses in contraceptive studies [12], and other terminology harmonization efforts, especially those using methodology to establish consensus and priorities (e.g., Delphi method, Child Health and Nutrition Research Initiative [CHNRI] approach).

<sup>iii</sup> Definition of instrument: An approach or method to measure or assess the outcome of interest, including those that are: (a) participant/patient-reported, (b) researcher/clinician-reported, (c) observer-reported; (d) biomarker-based, or (e) assay-based, as well as both ‘objective’ or ‘subjective’ measurements that encompass both direct observations and/or perceptions (Adapted from [56])

<p>1.3 * †</p>	<p>Additional research will be needed for alignment with recommendations 1.1, 1.2, and 2.1 to 2.3.</p> <p>Presently—and until such research is complete—researchers should use the following standardized definitions for spotting and bleeding:</p> <ul style="list-style-type: none"> <li>a. <b>Spotting:</b> Blood and other uterocervical fluid for which the person uses no menstrual products.<sup>iv</sup></li> <li>b. <b>Bleeding:</b> Any amount of blood and other uterocervical fluid greater than spotting requiring menstrual product use.</li> </ul> <p>Researchers should continue to study how diverse populations understand and differentiate spotting and light bleeding to generate more standardized language for measurement across trial populations. <i>(Also see 2.3b.)</i></p> <p><i>(See 5.2 for additional definitions).</i></p>
<p>1.4</p>	<p>While it is important to keep CIMC measures consistent for comparability across trials, there should be some flexibility for local research teams to translate, adapt, and add to these tools based on the local culture and context.</p>
<p><b>2. Instruments for measuring CIMCs</b></p>	
<p>2.1 * §</p>	<p>Researchers will need to develop new, standard instruments to measure CIMCs in trials because no validated instrument currently exists to address CIMC constructs of interest and for use in the context of contraceptive clinical trials [43]. Instruments for both (a) the menstrual changes themselves (e.g., changes in bleeding frequency, duration, volume; change in blood; changes in pain) and (b) the impact, acceptability<sup>v</sup>, and perceptions of those menstrual changes will be needed. Development of new, standard instruments can include: (a) generating and validating new items, and/or (b) modifying items from related, existing instruments used to measure changes to the menstrual cycle (e.g., instruments developed within the context of examining menstrual &amp; gynecological disorders &amp; symptoms).</p>

<sup>iv</sup> In some populations, the availability of menstrual products may be limited and thus reduce their use. Other populations may readily access menstrual products and use them differently than other populations would. When a trial is performed in such a population, the authors should describe this issue and how they choose to define spotting and bleeding for their particular population.

<sup>v</sup> Acceptability [of CIMCs] refers to a multidimensional concept wherein a user considers the CIMCs they experience to be suitable based on their individual circumstances. When considering whether a CIMC is acceptable, a user may evaluate their satisfaction and comfort with the CIMC, feelings regarding the burden/impact related to the CIMC, and the extent to which the CIMC fits within their value system [4,57]. By characterizing a CIMC as acceptable, a user effectively considers the CIMC to be suitable based on their individual anticipated, perceived, or experiential responses [7]. Acceptability will be influenced by a user's own unique attitudes and context. It is shaped by individual and community-level norms, preferences, and practices regarding menstrual health management and self-efficacy to manage the CIMC.

	<p>Because additional research will be needed to develop and validate instruments in alignment with this recommendation, 4.1 to 4.13 describe data collection approaches recommended for trials conducted prior to this additional research and 5.1 to 5.13 describe related data analysis approaches.</p> <p>Recommendations 2.2 and 2.3 provide additional details on instrument development and validation.</p>
2.2 ** §	<p>When developing instruments to measure CIMCs in trials, it is critical that researchers:</p> <ol style="list-style-type: none"> <li>Follow current recommendations from regulatory agencies and professional societies and consortia, such as the FDA’s guidance on PRO<sup>vi</sup> use in trials [26] and the PROTEUS Consortium’s recommendations on measuring PROs effectively.<sup>vii</sup></li> <li>Use simple, evidence-based patient-centered terminology both within instrument items, prioritizing the perspective of the contraceptive user and simplicity over the convenience of the researcher or imposing definitions or characterizations that may not be sufficiently objective to translate across contexts.<sup>viii</sup></li> <li>Ensure instruments are accessible to those with differing literacy levels and people with disabilities, including consideration of pictorial instruments and adherence to current USG plain language guidelines [44] and Web Content Accessibility Guidelines (WCAG) [45] for digital content.</li> <li>Ensure instruments adhere to data privacy requirements, legal protections, and ethical considerations to protect data and prevent use of data by any entity outside of those described in the trial informed consent forms.</li> </ol>
2.3 §	<p>When developing instruments to measure CIMCs in trials, ideally:</p> <ol style="list-style-type: none"> <li>For participant-entered diary instruments, researchers should establish evidence<sup>ix</sup> on recall to determine the window of retrospective data entry permitted for each type of CIMC measured; such evidence may build upon previous research around heavy menstrual bleeding (HMB).</li> <li>Researchers should establish the transferability<sup>x</sup> of instruments, including following recommendations such as the ISPOR principles on good practice for translation and cultural adaptation [46,47] because instruments are needed in different languages and sociocultural contexts, including for multisite trials. When feasible, this should include simultaneous development and validation of instruments in multiple contexts. (<i>Also see 1.3.</i>)</li> </ol>

<sup>vi</sup> All data on CIMCs in trials are reported by the participant (i.e., it is a patient reported outcome [PRO]) unless the trial is using an indirect quantitative and semi-quantitative approach like an alkaline hematin assay to determine menstrual blood loss from used menstrual products. See 4.6, 5.7b, 6.1, and 6.2 for recommendations on these types of instruments.

<sup>vii</sup> PROTEUS recommendations on measuring PROs effectively with ISOQOL standards [58].

<sup>viii</sup> If supported by evidence, an example of this could be asking participants about use/nonuse of menstrual products, instead of asking participants to use and understand the term ‘spotting’, which may not be relevant in all contexts.

<sup>ix</sup> “Establishing evidence” can include reviewing existing research and/or conducting new research.

<sup>x</sup> Transferability is the degree to which an instrument can be transferred between linguistic and sociocultural groups.

	<ul style="list-style-type: none"> <li>c. Researchers should consider time and cost requirements to implement instruments, so they do not unduly hinder use in contexts with fewer resources.</li> <li>d. For mixed mode instruments (e.g., paper and electronic instruments, or interactive voice response (IVR) and short message service (SMS)/chat instruments), researchers should establish equivalence<sup>xi</sup> between modes (e.g., between paper and electronic versions of an instrument), including following recommendation such as the ISPOR good research practices on use of mixed mode PROs [48].</li> <li>e. For instruments requiring data entry by participants on electronic hand-held devices (i.e., phones or tablets), researchers should establish equivalence<sup>xi</sup> for devices provided to participants by the trial as well as for devices already used/owned by participants (i.e., BYOD, ‘bring your own device’) so either can be used by trials without the potential of introducing differential bias between trials providing devices and trials with BYOD or within trials where some participants are provided devices and some are BYOD [49].</li> <li>f. Researchers should consider if additional instrument items or modules are needed for participants from populations who may have different CIMC outcomes and experiences (e.g., those with menstrual &amp; gynecological disorders &amp; symptoms).</li> </ul>
<b>3. Trial design, protocol development, participant recruitment</b>	
3.1	Early phase studies should consider evaluation of drug-drug interactions—including pharmacokinetic (PK) or pharmacodynamic (PD) interactions that may impact CIMCs—for commonly used medications that may influence CIMCs. These medications can include over the counter (OTC), supplements, or even certain foods.
3.2 *	Every effort should be made to evaluate a range of body mass index (BMI) <sup>xii</sup> categories in PK studies, as BMI heterogeneity can have major implications for PK and/or CIMCs.

<sup>xi</sup> Measurement equivalence is a function of the comparability of the psychometric properties of the data obtained via the original and adapted administration mode... to demonstrate that the change did not introduce response bias and/or negatively affect the measure’s psychometric properties” [48].

<sup>xii</sup> There are many critiques of BMI as an indicator and important concerns about its history and use [59–62]; however, we also note current regulatory use of BMI, including in FDA requirements for hormonal contraceptive clinical trials (i.e., “The effectiveness of some contraceptives may be reduced with increasing body weight. Sponsors should not place restrictions on body mass index (BMI) for trial enrollment. The trial population should include obese women (i.e., defined as BMI of at least 30 kg/m<sup>2</sup>), and the analysis plan should include a prespecified subgroup efficacy analysis in this population. Insufficient data in the obese population may result in a limitation of use for this population in labeling. During the trial design phase, sponsors should discuss with the division the adequacy of the number of cycles of drug exposure that will be derived from obese subjects.”; and “For the overall trial population, as well as for those younger than or equal to 35 years old at study enrollment and those older than 35 years old, sponsors should perform subgroup analyses [of efficacy] based on BMI at study enrollment...”) [63].

3.3 *	<p>Related to intrinsic or extrinsic factors that may impact CIMCs:</p> <ul style="list-style-type: none"> <li>a. Researchers should consider the most relevant intrinsic or extrinsic factors and design enrollment for Phase III and IV trials to include adequate numbers of participants with these factors. The choice of relevant factors will be informed by trial context, trial population, product characteristics, and/or trial design.</li> <li>b. Information on intrinsic and extrinsic factors should be collected during the trial, as well as at enrollment.</li> <li>c. When the primary trial outcome is pregnancy prevention, researchers should consider the impact of the selected intrinsic and extrinsic factors on risk of pregnancy.<sup>xiii</sup></li> </ul>
3.4 §	<p>Researchers should establish and standardize what data are needed at enrollment on menstrual cycle history (when not using hormonal or intrauterine contraception) and other sexual and reproductive health history based on the CIMC data collected during the trial in order to permit comparisons between baseline and treatment outcomes. Additional baseline data on sociodemographics and other relevant background data may also be needed for stratified analyses to examine relevant intrinsic and extrinsic factors that may impact CIMCs.</p>
3.5	<p>When feasible, in Phase II and III trials, measurement of CIMCs and related outcomes should include input from local researchers, and whenever possible, also be under the advisement of local community advisory boards and community advocates. This input should incorporate input from a variety of stakeholders, including partners and families.</p>
3.6	<p>In Phase III and IV studies, diverse trial sites should be used to more fully represent differences in CIMC experiences and sociocultural differences in menstruation perceptions/knowledge across populations of future users.<sup>xiv</sup></p>
3.7 *	<p>Phase IV studies should be designed to build upon Phase III findings, to collect data on additional CIMC-related factors that may influence tolerability and method continuation, and to include adequate numbers of participants with these factors. The factors of interest may differ by trial context, product characteristics, and/or trial design.</p>

<sup>xiii</sup> Researchers may consider the factor of age as an example for how these relevant intrinsic and extrinsic factors that may impact risk of pregnancy could be addressed in trials (i.e., FDA requirements for hormonal contraceptive clinical trials: “Trials should include subjects from all premenopausal age groups who are likely to use the drug product, including postmenarchal adolescents.”; “The primary efficacy results should be calculated using the trial population of women younger than or equal to 35 years old at study enrollment because the likelihood of pregnancy decreases with advancing age. Include additional efficacy analyses for the overall trial population and a subgroup analysis for those older than 35 years old.” and “Enrollment of subjects older than 35 years old is recommended for safety determinations. The number of subjects older than 35 years old who should be enrolled in the trial or trials will depend on the existing experience with the drug product ingredients and should be discussed with the division [of Bone, Reproductive, and Urologic Products in the Center for Drug Evaluation and Research at the Food and Drug Administration]” [63].

<sup>xiv</sup> This site diversity may go beyond Stringent Regulatory Authorities [SRA] minimums and should include recruiting and enrolling participants within and between trial sites from different racial and ethnic groups, socioeconomic statuses, age categories, disability status, sexual orientation, and gender identity, as these all may influence CIMC experiences and perceptions of menstruation.



3.8	All health care providers who will be involved in contraceptive clinical trials should be trained in comprehensive counseling methods, including providing information about different CIMCs and management strategies for CIMCs that are culturally and contextually appropriate and referring to appropriate services when indicated.
<b>4. Data collection</b>	
4.1 **	<p>Additional research will be needed to align with recommendations 2.1 to 2.3. For trials conducted prior to this additional research, recommended data collection approaches are presented in 4.1 to 4.13 and recommended analysis approaches for these data are presented in 5.1 to 5.12.</p> <p>Presently, researchers should include two types of assessments for primary outcomes of CIMCs:</p> <ol style="list-style-type: none"> <li>a. Bleeding and spotting days, and</li> <li>b. Frequency, duration, and volume of bleeding episodes.</li> </ol>
4.2 §	<p>Presently, researchers may also collect data on additional PROs such as cramping, pain, changes in blood and other uterocervical fluid, as well as other symptoms that have been shown to influence patient experiences with and perceptions of contraceptive methods. <i>See recommendations 4.11 to 4.13 for details.</i></p> <p>More research is needed to create validated instruments for these types of additional outcomes across trial contexts. <i>See 2.1 to 2.3 for details on instrument development and validation recommendations.</i></p>
4.3 *	<p>Presently, researchers should collect <i>bleeding and spotting data</i> in trial phases as follows:</p> <ol style="list-style-type: none"> <li>a. In Phase I and IIa trials, researchers may choose to capture bleeding and spotting assessments as an exploratory outcome to make early decisions about product viability.</li> <li>b. In Phases IIb and III, researchers should collect bleeding and spotting data.</li> <li>c. In Phase IV, bleeding and spotting assessments may be required to assess additional use/benefits of the contraceptive (e.g., for special populations, health conditions, comparisons with other contraceptives).</li> </ol>
4.4 *	<p>Presently, researchers should collect data on <i>frequency, duration, and volume</i> of bleeding episodes in trial phases as follows:</p> <ol style="list-style-type: none"> <li>a. In phases IIb and III, these assessments are recommended.</li> <li>b. In phase IV, these assessments may be required for additional use/benefits of the contraceptive (e.g., for special populations, health conditions, comparisons with other contraceptives).</li> </ol>
4.5	Presently, researchers should collect some bleeding data for the entire duration of the method, as practical, including for methods with longer durations (e.g., 3-5 years).
4.6 *	Presently, researchers should use validated measures for an objective measure of <i>bleeding volume</i> in trials when bleeding volume is an important outcome for label purposes (e.g., as an indication for HMB).
4.7	Presently, researchers should design bleeding data capture instruments to register information daily when possible and, if not, at least each time when bleeding or spotting occurs. However, researchers should study alternatives to daily collection of data for methods that have long durations (e.g., 3-5 years).

	<i>(See 2.3a on recall for new instruments.)</i>
4.8	Presently, when using electronic diaries, the diaries should be designed to allow a window of time for users to retroactively record bleeding occurrences (e.g., within 48 hours); however, more research needs to be conducted to find the optimal window of time. <i>(See 2.3a on recall for new instruments.)</i>
4.9 *	Presently, researchers should consider the culture and context of their trial setting to determine whether paper or electronic diaries will be more acceptable and/or feasible and assess the strengths and limitations of each method of data collection. <i>(See 2.3d on mixed modes for new instruments.)</i>
4.10	Presently, researchers using electronic diaries should attempt to incorporate data capture on a participant’s own phone to avoid the burden of using additional devices where possible, despite software standardizing challenges. <i>(See 2.3e on BYOD for new instruments.)</i>
4.11 ** §	Presently, trials should measure acceptability and quality of life related to CIMCs to assess the impacts of CIMCs on users and inform contraceptive counseling and user choices. This can be done in the following ways: <ul style="list-style-type: none"> <li>a. In Phase I, II, and III, researchers should measure participants’ self-perceived changes in bleeding and other CIMCs (e.g., whether participants self-define changes as significant). This should be measured over time for Phase II and III.</li> <li>b. In Phase II and III, researchers should measure the physical and psychosocial impacts of CIMCs on quality of life, including burdens and benefits. The aspects that are measured should be based on local priorities and informed by local researchers and advisors, and may include impacts on work, sexual well-being, mental burden, anxiety and stress, quality of relationships, social well-being, and financial well-being (including the burden of menstrual health management).</li> <li>c. In Phase II and III, researchers should measure changes in acceptability related to CIMCs over time.</li> </ul>
4.12	Presently, in Phase III, and when possible in Phase II, trials should include measures to contextualize and understand the CIMC acceptability and quality of life impacts outlined in recommendation 4.11. This can be done in the following ways: <ul style="list-style-type: none"> <li>a. Trials should collect sufficient demographic data, including about race and ethnicity, socioeconomic status, age, disability, sexual orientation, gender identity, location (e.g., urban/rural), and other social factors.</li> <li>b. Trials should collect data about affordability and access to menstrual health information, menstrual products/materials, and water and facilities for changing, washing, and disposing of menstrual products.</li> <li>c. Trials should collect data about the ethicality and cultural appropriateness of CIMCs (e.g., shame or stigma surrounding CIMCs) in the trial context, along with individuals’ beliefs about CIMCs (e.g., the relationship between CIMCs and long-term health outcomes like infertility or cancer).</li> <li>d. Researchers should measure self-efficacy related to managing CIMCs, that is, an individual’s confidence in their capacity to manage the CIMCs experienced.</li> </ul>
4.13	Presently, in all phases of trials, acceptability and impacts of CIMCs should be measured using both quantitative measures and qualitative methods whenever possible and advantageous. Qualitative methods will vary based on context and type of

	trial but may include open-ended survey questions, in-depth interviews, or ranking exercises. Research teams should include members experienced in qualitative measurement and analysis whenever these methods are being used.
<b>5. Data analysis</b>	
5.1 †	<p>Additional research will be needed to align with recommendations 1.1, 1.2 and 2.1 to 2.3, which will include establishing how new instrument(s) will be analyzed. For trials conducted prior to this additional research, recommended data analysis approaches are presented in 5.1 to 5.12 for data collected per 4.1 to 4.13.</p> <p>Presently, to align with a patient-centered approach, data in the recommendations below should be reported for each 30-day <b>reference period</b>. To permit comparison to existing methods, data should also be reported for each 90-day reference period as well as other reference period length if relevant for the method.<sup>xv</sup></p>
5.2 ** †	<p>Presently, researchers should use the following definitions during analysis for bleeding episode and for duration, frequency, and volume of bleeding episodes:</p> <ol style="list-style-type: none"> <li>a. <b>Bleeding episode:</b> one or more consecutive days with bleeding and/or spotting, bordered by 2 full days without any bleeding or spotting.</li> <li>b. <b>Duration:</b> A bleeding episode can be classified into prolonged or not prolonged. <ul style="list-style-type: none"> <li>• <u>Prolonged Duration</u> – Bleeding/spotting episode lasting more than 7 days.</li> </ul> </li> <li>c. <b>Volume:</b> Volume should be classified according to the participant’s perception. Volume should be described as “lighter,” “usual,” or “heavier” as compared to the individual’s typical volume when not using contraception.</li> <li>d. <b>Frequency:</b><sup>xvi</sup> <ul style="list-style-type: none"> <li>• <u>Absence of bleeding/spotting:</u> no bleeding or spotting during the reference period.</li> <li>• <u>Infrequent bleeding:</u> 2 or fewer bleeding/spotting episodes during a 90-day reference period.</li> <li>• <u>Frequent bleeding:</u> &gt; 4 bleeding/spotting episodes during a 90-day reference period.</li> </ul> </li> </ol>

<sup>xv</sup> For example, 28 days for a combine oral contraceptive with 24/4 regimen (i.e., 24 days of pills with the APIs and 4 days of placebo pills) or other single cycle regimens (e.g., 21/7); or 84 days for an extended cycle combined oral contraceptive [12].

<sup>xvi</sup> We have not included a “normal” frequency definition. Earlier versions of this recommendation included a definition for “approximately monthly bleeding” as “3 or 4 bleeding/spotting episodes during the 90-day reference period”, similar to that proposed by Creinin et al. [12]. Without clarification on the duration between these episodes, however, some experts raised the concern that a bleeding pattern could fit this definition but not actually be “monthly” or “normal” (e.g., 3 or 4 episodes in the first 30 days and then none in the latter 60 days). See 5.5a for an alternative, patience-centered frequency definition. Further work towards consensus on an appropriate definition should be part of future updates to these recommendations.

	<ul style="list-style-type: none"> <li>• <u>Scheduled bleeding</u>: a bleeding/spotting episode starting during the expected interval and lasting no more than 7 days for a method expected to have a predictable bleeding pattern.<sup>xvii</sup></li> <li>• <u>Unscheduled bleeding</u> a bleeding/spotting episode occurring at any time prior to the expected interval, even if it continues into the next expected interval for a method expected to have a predictable bleeding pattern.<sup>xvii</sup></li> </ul>
5.3 **	<p>Presently, analysts should report the following <b>variables</b> on <i>number of bleeding and spotting days</i> during the reference period:</p> <ol style="list-style-type: none"> <li>Total number of bleeding or spotting days,</li> <li>Number of bleeding days,</li> <li>Number of spotting days,</li> <li>Number of scheduled and unscheduled bleeding or spotting days, if relevant for the method as described in 5.2d.</li> </ol>
5.4 *	<p>Presently, analysts should report the following <b>descriptive statistics</b> on <i>number of bleeding and spotting days</i> during the reference period:</p> <ol style="list-style-type: none"> <li>For Phase I and IIa trials, as exploratory summary statistics: <ul style="list-style-type: none"> <li>• Mean, median, interquartile range (IQR) for: <ul style="list-style-type: none"> <li>○ Bleeding or spotting days,</li> <li>○ Bleeding days,</li> <li>○ Spotting days,</li> <li>○ Instances of consecutive days with no bleeding if daily bleeding data were collected, and</li> <li>○ Scheduled and unscheduled bleeding or spotting days, if relevant for the method as described in 5.2d.</li> </ul> </li> <li>• Percent of participants with no bleeding or spotting.</li> </ul> </li> <li>For Phase IIb, III and IV trials: <ul style="list-style-type: none"> <li>• Mean, median, range, interquartile range (IQR),</li> <li>• An estimate of precision (e.g., standard error of the mean or confidence interval)</li> <li>• These data should be presented separately for total, unscheduled, and scheduled days if relevant for the method as described in 5.2d.</li> </ul> <p>However, analysts are advised to consider whether the sample size and presumed distribution of the data warrants the use of standard statistical methods.</p> </li> </ol>
5.5	<p>Presently, analysts should report the following <b>variables</b> on <i>bleeding frequency</i> during the reference period.</p>

<sup>xvii</sup> Predictable bleeding pattern: an expected predictable pattern is based on the product design as determined by the regimen (e.g., combined hormonal methods [oral, transdermal, vaginal, or injectable] and non-hormonal methods such as IUDs or non-hormonal vaginal products). Unpredictable bleeding pattern: an expected unpredictable pattern based on the product design (e.g., progestin only methods including implants, IUDs, injectables, pills, patches, and rings) [12].

	<p>a. To align with a patient-centered approach (See 5.1):</p> <ul style="list-style-type: none"> <li>• Whether bleeding episodes occur (1) More than once per month, (2) Once per month, or (3) There is a month without bleeding</li> <li>• Whether there were no bleeding episodes for three or more months in a row, if relevant for the method</li> </ul> <p>b. To permit comparison to data from existing methods using a 90-day reference period:<sup>xvi</sup></p> <ul style="list-style-type: none"> <li>• No bleeding or spotting</li> <li>• Infrequent bleeding</li> <li>• Frequent bleeding</li> <li>• Scheduled and unscheduled bleeding or spotting, if relevant for the method as described in 5.2d</li> </ul>
5.6	<p>Presently, analysts should report the following <b>variables</b> on <i>bleeding duration</i> during the reference period.</p> <p>a. To align with a patient-centered approach (<i>see 5.1</i>): Mean, median, interquartile range (IQR) for number of bleeding days per bleeding episode.</p> <p>b. To permit comparison to data from existing methods using a 90-day reference period: The number of participants with prolonged bleeding.</p> <p>c. Bleeding duration variables should be reported separately for scheduled and unscheduled bleeding episodes, if relevant for the method as described in 5.2d.</p>
5.7	<p>Presently, analysts should report the following <b>variables</b> on <i>bleeding volume</i> during the reference period.</p> <p>a. To align with a patient-centered approach (<i>see 5.1</i>): An assessment by the participant per 5.2c</p> <p>b. If indirect quantitative or semi-quantitative measurement of volume of bleeding is needed for the trial, analysis will depend on how data are collected (i.e., alkaline hematin assay, PBAC score; <i>see 4.6</i>)</p> <p>c. Bleeding volume variables should be reported separately for scheduled and unscheduled bleeding episodes, if relevant for the method as described in 5.2d.</p>
5.8	<p>Presently analysts should report the following <b>descriptive statistics</b> on <i>bleeding frequency</i> and <i>bleeding duration</i> during the reference period:</p> <p>a. For Phase I and IIa trials: same as Phase IIb, III, and IV trials when the trial is longer than 90 days.</p> <p>b. For Phase IIb, III, and IV trials: Absolute and relative frequencies of the variables for frequency and duration of bleeding episodes</p>
5.9	<p>Presently, analysts should report the following <b>inferential statistics</b> on number of <i>bleeding and spotting days</i>, as well as <i>frequency, duration and volume of bleeding</i>, during the reference period:</p> <p>a. For Phase I and IIa trials: not include any formal hypothesis testing, but exploratory analysis could be used to exclude contraceptives with undesirable CIMCs.</p> <p>b. For Phase IIb, III and IV trials: Use hypothesis testing for studies with comparator/control and use an analysis method that accounts for repeated measures on subjects, such as generalized estimating equations (GEE).</p>

5.10 *	Presently, for <i>Phase IIb, III, and IV trials</i> , analysts should report <b>sensitivity analysis</b> to evaluate the effect of some variables of interest on bleeding and spotting (e.g., age, body mass index, smoking status). If feasible and informative, sensitivity analyses can be done for <i>Phase I and IIa trials</i> .
5.11 **	Presently, analysts should report the following about <b>missing data</b> <sup>xviii</sup> on the number of <i>bleeding and spotting days</i> , as well as <i>frequency, duration and volume of bleeding</i> , during the reference period: <ul style="list-style-type: none"> <li>a. For Phase I and IIa trials: The number of discontinuations and the reasons for discontinuation</li> <li>b. For Phase IIb, III and IV trials: The number of discontinuations and the reasons for discontinuation; and the potential impact of informative drop-out when describing CIMCs over time if a meaningful percentage of trial participants discontinue due to CIMCs<sup>xix</sup></li> </ul>
5.12	Presently, analysts should include <b>data visualizations</b> , such as heatmaps and boxplots, for data from all participants for all trials.
5.13	Data on CIMC acceptability should be analyzed as dictated by the instrument used for measurement. Qualitative data should be analyzed via standard qualitative methods appropriate for the qualitative approach.
5.14 *	Following analysis, researchers/analysts should (a) deposit deidentified data in an established open data repository, as soon as possible; (b) deposit deidentified CIMC data from assays in an established open data repository; and (c) aim to make instruments openly available for broad use, ideally at no cost. These steps should be done whenever possible, such that the community can benefit from shared knowledge and minimized duplication of resources. <sup>xx</sup>
<b>6. Areas for exploratory research<sup>xxi</sup></b>	
6.1	When developing non-PRO instruments to measure CIMCs in trials to advance research on biospecimen collection and assays: Researchers should establish a target assay profile <sup>xxii</sup> for instruments used by trial investigators and/or participants for

<sup>xviii</sup> Data may be missing due to early method discontinuation, trial discontinuation, or loss to follow up.

<sup>xix</sup> This is to account for the distribution of CIMC outcomes towards the end of the trial being disproportionately represented by subjects who did not experience those outcomes or experienced less impactful outcomes. Approaches could include (a) reporting CIMC outcomes separately among all enrolled participants (where the denominator is the number of subjects still in follow-up at a given time) and among participants who completed the trial (where the denominator is uniform over time), where differences in patterns would suggest that informative drop-out was biasing the results; or (b) explicitly modeling the (potentially informative) drop-out mechanism using an appropriate methodology, such as the one developed and performed in the context of DMPA-induced amenorrhea [64,65] using weighted generalized estimation equations (GEE) for population-level analyses [66].

<sup>xx</sup> There are notable concerns about the extent to which data can be truly deidentified, especially as larger amounts of open data continue to become available [67–69].

<sup>xxi</sup> For additional CIMC research, see the Global CIMC Task Force research and learning agenda on CIMC measurement broadly, biomedical research, social-behavioral and user preferences research, and programmatic research [1].

<sup>xxii</sup> This can be analogous to a drug/device target product profile (TPP) with minimum acceptable result and ideal results (e.g., [70]).

	<p>biospecimen collection and assays. Example ideal characteristic criteria could be affordability, acceptability, ease of storage without cold chain requirement, low burden on participants and/or trial investigators (e.g., non-invasive, ease of collection/assay), speed of results, and ease of interpretability. Urine pregnancy tests can be used as a model for biospecimen collection and assays.</p>
6.2	<p>To advance a future research agenda for measurement of CIMCs in trials, researchers should explore avenues for developing new or improved assays and identifying biomarkers<sup>xxiii</sup> or other approaches such as predictive modeling for CIMCs. However, researchers should always prioritize protecting participant privacy and data security when considering the use of new methodology, especially when using Artificial Intelligence or Big Data approaches [50–54].</p>

---

<sup>xxiii</sup> Developing assays and biomarkers for trials should be done in consultation with regulatory agencies.