Contraception 112 (2022) 14–22

Contents lists available at ScienceDirect

Contraception

journal homepage: www.elsevier.com/locate/contraception

Commentary

Recommendations for standardization of bleeding data analyses in contraceptive studies

Mitchell D. Creinin[a,*], Carolina Sales Vieira[b], Carolyn L. Westhoff[c], Diana J.A. Mansour[d]

[a] Department of Obstetrics and Gynecology, University of California, Davis, Sacramento, CA, United States
[b] Department of Obstetrics and Gynecology, Ribeirao Preto Medical School, University of São Paulo, São Paulo, Brazil
[c] Department of Obstetrics and Gynecology, Columbia University, New York, NY, United States
[d] Integrated Sexual Health Service, Newcastle upon Tyne Hospitals NHS Foundation Trust, United Kingdom

ARTICLE INFO

Article history:
Received 28 January 2022
Received in revised form 20 May 2022
Accepted 24 May 2022

Keywords:
Contraception
Bleeding
Pattern
Flow
Duration
Recommendations

ABSTRACT

To address limitations that exist with existing definitions of menstrual bleeding changes that occur with contraceptive methods, we assembled a panel to develop new recommendations for standardization of bleeding data analyses associated with contraceptive use to better inform users, clinicians, investigators, pharmaceutical companies, and regulatory agencies. We propose three criteria for assessing bleeding outcomes: pattern, flow, and duration. The descriptors within each criterion depend on whether the contraceptive is designed to result in a predictable or unpredictable bleeding pattern. Predictable pattern outcomes quantify days of scheduled, unscheduled and no bleeding, while unpredictable pattern outcomes assess frequency. Flow is quantified based on patient comparisons to their typical flow when not using contraception, with spotting representing no menstrual products use. Duration of a prolonged bleeding and/or spotting episode is more than 7 days. Studies should assess bleeding characteristics for a minimum of 12 months for 21/7, 24/4, extended cycle or continuous regimens, two years for injectables, and the full duration of use for long-acting contraceptives. Describing pattern, flow and duration as independent categories allows a fuller understanding of the bleeding outcomes and better future assessments of acceptability and continuation. Standardization of outcomes permits better comparison between studies and data synthesis; standardization will also improve the ability of clinicians and patients to understand differences between products.

1. Introduction

Although contraceptives are approved by regulatory agencies based primarily on efficacy and safety, secondary outcomes, particularly changes in bleeding, influence contraceptive choice, user satisfaction, and continuation. Some societies, cultures, and individuals are less prone to accept menstrual changes; thus, bleeding complaints are one of the most common reasons for discontinuation [1–5]. Decrease in flow has led to wide acceptance of certain methods, such as the levonorgestrel 52 mg intrauterine device [6,7], yet less flow or amenorrhea may be poorly accepted, especially if unexpected [8,9].

In 1986, the World Health Organization (WHO) proposed standardized terminology for and analysis of contraceptive-induced bleeding changes, known as the Belsey criteria [10]. These criteria used 90-day segments (called "reference periods") as first suggested by Rodriguez et al. in 1976 [11], with the goal that this time frame would be long enough to allow an overall description of the flow and frequency of bleeding episodes (Table 1). The definitions described flow as bleeding or spotting episodes that may or may not require menstrual products. Within each 90-day reference period, the authors recommended dividing the patterns of these episodes into amenorrhea or 1 of 4 specific categories defined as prolonged, frequent, infrequent, or irregular bleeding [10]. Belsey and colleagues [12] argued that 1 month or cycle is too short for "…assessments of events which occur at intervals often exceeding 30 days and does not allow reliable estimation of the average length and variability of events for the individual subjects."

In 2005, a consensus panel convened to address apparent deficiencies in bleeding analyses in contraceptive studies [13]. The Belsey criteria were not being applied routinely in clinical trials. More importantly, because the criteria were not designed to describe planned cyclic bleeding, they did not apply well to combined hormonal contraceptives. Pharmaceutical companies and

* Corresponding author.
E-mail address: mdcreinin@ucdavis.edu (M.D. Creinin).

https://doi.org/10.1016/j.contraception.2022.05.011
0010-7824/

Table 1
World Health Organization Belsey definitions of bleeding patterns with contraceptive use [10]

Term	Definition
Amenorrhea	No bleeding or spotting during a 90-d reference period
Prolonged bleeding	Bleeding/spotting episodes lasting more than 14 d during a 90-d reference period
Frequent bleeding	More than 5 bleeding/spotting episodes during a 90-d reference period
Infrequent bleeding	One or 2 bleeding/spotting episodes during a 90-d reference period
Irregular bleeding	Three to 5 bleeding/spotting episodes and less than 3 bleeding/spotting-free intervals of 14 d or more during a 90-d reference period

regulatory agencies generally did not follow these guidelines, and pharmaceutical companies often developed novel bleeding categories to maximize the positive aspects of their results. The 2005 panel provided recommendations based on best practice in trial design, data collection, and analysis focusing on combined hormonal methods, creating what is known today as the "Mishell" criteria [14].

More than thirty-five years since the Belsey criteria and 15 years since the Mishell criteria, we face the same difficulties. The Belsey criteria are commonly used for injectable and long-acting methods, and the Mishell criteria, when used, are applied to short-acting methods such as pills, rings, and patches [10–14]. However, not all contemporary short-acting methods are designed to have predictable monthly bleeding. Whereas cyclic methods were most consistently used in a 21-day hormone/7-day hormone-free interval pattern decades ago, many formulations are now 24/4 or have extended use for 3 months or more. Several studies serve as examples of how contraceptive-induced menstrual bleeding changes are reported heterogeneously, even when evaluating the same product [15–17]. The lack of standardization among studies assessing these changes underscores that the current recommendations to describe the changes are inadequate [4].

The authors have worked in this topic area over the last 2 or more decades. Three of the authors (MDC, CSV, DM) recognized many of the issues during a recent collaborative analysis and invited the fourth author to join the panel (CLW) based on expertise. The panel had initial conversations in May 2020 followed by a general outline of a proposal for funders to obtain support for time to develop the recommendations and publication-related costs. We obtained funding as unrestricted grants or gifts from multiple international pharmaceutical companies that manufacture contraceptives, expecting they would support use of the final output. We initiated meetings and writing of draft recommendations in May 2021. After follow-up meetings and correspondence to reach a consensus, we completed an initial set of recommendations in January 2022. After review by the Society of Family Planning Clinical Affairs Committee and peer review, this document represents our final recommendations.

Bleeding outcome assessments include a description of contraceptive-induced bleeding, acceptability, and therapeutic use in patients with heavy bleeding complaints. Our panel aimed to create contemporary recommendations specific to assessments of contraceptive-induced bleeding with the use of modern hormonal and nonhormonal contraceptives. We recognized that the Belsey and Mishell criteria were developed by and for researchers without consideration of how patients might view bleeding descriptors. While investigators have evaluated bleeding outcomes with contraceptive use and patient acceptance of these bleeding profiles, we found no research that focused on the descriptors. Both sets of previous criteria used bleeding definitions that failed to acknowledge a patient's subjective view; for example, one person may prefer to use a panty liner when spotting while someone else may not.

We developed these recommendations for contraceptive users, clinicians, investigators, pharmaceutical companies, and regulatory agencies. In addition, these recommendations are targeted to allow clinicians and patients to understand contraceptive-induced bleeding in a simpler way, to aid in counseling and comparison between products.

Our goal with these recommendations is to allow future studies to describe bleeding characteristics in a specific manner that facilitates clinicians' and patients' ability to make informed choices. Daily data collection is still needed for any analyses, so existing data sets from past studies will be amenable to most, but not all, of this proposed framework. Standardization of outcomes will allow better comparison between studies and data synthesis across studies in line with the Core Outcomes in Effectiveness Trials (COMET) initiative [18,19]. Additionally, standardization will help inform communication among those working to manage bleeding complaints.

2. Criteria consideration

We assume that contraceptive users do not want their bleeding outcomes to be worse than their norm. Some may be willing to accept a worse bleeding outcome as a trade-off for other positive features of a contraceptive, but we believe this is unlikely a preferred result. Typically, individuals want bleeding to be the same or better; for some, less or no bleeding may be considered an improvement. Importantly, everyone's baseline is different, and their individual assessment of baseline bleeding is not uniform. For example, amongst those with heavy menstrual bleeding (blood loss >80 mL based on alkaline hematin testing), 41% considered their flow moderate or scanty while 14% with blood loss <20 mL considered their flow heavy [20].

Using consistent terminology across studies to describe bleeding outcomes is key. Confusion occurs when investigators or authors define terms differently. For decades, 'bleeding patterns' has commonly been used to describe all of the criteria related to contraceptive-induced bleeding. Here, we attempt to outline clearer terminology that differentiates 3 unique outcome criteria: pattern (referring specifically to when bleeding occurs or does not occur, i.e., frequency), flow, and duration.

The Belsey criteria include several terms used as pattern descriptors: amenorrhea, prolonged bleeding, frequent and infrequent bleeding, and irregular bleeding (Table 1). These pattern definitions are not mutually exclusive – e.g., a user could experience infrequent, prolonged episodes; yet, authors have erroneously analyzed bleeding outcomes as if the definitions are exclusive [17]. Thus, the description needs to separate characteristics of the bleeding episode based on frequency (frequent, normal frequency, or infrequent) and duration (prolonged or not prolonged). The Belsey definition included irregular bleeding and did not define regular. However, because regularity (and by default, irregularity) of bleeding is a measure of frequency, the use of regular or irregular as a pattern is duplicative and is not included in these new recommendations.

The Belsey definition of a prolonged bleeding episode (more than 14 days) seems excessively long, as users would not be likely to speak favorably about 14-day bleeding episodes, but negatively about 15-day episodes. A 2019 analysis of bleeding patterns with

the etonogestrel implant found no difference in discontinuation rates with a definition of prolonged bleeding shortened to more than 7 days [21]. In that study, investigators examined discontinuation rates among users with prolonged bleeding who had an otherwise favorable bleeding pattern (amenorrhea, infrequent bleeding, or normal frequency bleeding with no prolonged bleeding) in the first few months of use. Discontinuation rates did not differ with defining prolonged bleeding as more than 14 days (12/325 [3.7%]) or more than 7 days (5/211 [2.4%]), $p = 0.46$ (Fisher exact test). Of note, among implant users with an initial unfavorable pattern (prolonged or frequent bleeding), discontinuation rates with prolonged bleeding are defined as more than 14 days (27/212 [12.7%]) or more than 7 days (34/326 [10.4%]) were similar, $p = 0.41$. However, as the authors point out, the clinical trial data used in this assessment were more than 15 years old, and "more contemporary data may result in a different finding." We agree that a shorter definition is more appropriate and aligned with real-life data.

We evaluated other possible criteria published subsequent to the Belsey and Mishell guidelines that could be applied to new recommendations. The International Federation of Gynecology and Obstetrics (FIGO) created a classification system for causes of abnormal uterine bleeding in 2011 and updated the system in 2018 [22,23]. The authors defined abnormal uterine bleeding as "the symptom of disturbed menstrual bleeding" [22]. These criteria include frequency, regularity, duration, volume, and intermenstrual bleeding with the goal of defining normal and abnormal uterine bleeding in those not using contraception. The purpose of this FIGO classification is to consider when diagnostic testing or treatment may be indicated to identify pathologic conditions, which is different than our goal of defining bleeding outcomes for contraception users.

Another example is an attempt to simplify the description of the bleeding patterns. In a reanalysis of data from the phase 3 studies of the etonogestrel implant, Mansour et al. [20] classified the bleeding changes associated with this contraceptive into favorable (amenorrhea, infrequent bleeding, and normal frequency bleeding with no prolonged bleeding) or unfavorable (prolonged and/or frequent bleeding) based on the discontinuation rates associated with these patterns. Prolonged and/or frequent bleeding were associated with a higher etonogestrel implant discontinuation rate [21–24]. Which bleeding changes a contraceptive user will consider favorable/unfavorable or acceptable/not acceptable will depend on many factors such as age, culture, religion, life course, education, presence of symptoms associated with bleeding, and contraceptive options available [4]. A scoping review on responses to contraception-induced bleeding changes concluded that substantial variability exists in how contraceptive users respond, including what they prefer and what they are willing to tolerate in exchange for the contraceptive benefits. Importantly, these responses are shaped by individual and social influences. For example, stated preferences for absent bleeding ranged from 0% to 65% across studies included in the review [4]. Therefore, it is not possible to classify bleeding outcomes into positive/negative, acceptable/not acceptable, or favorable/unfavorable without knowing individual bleeding preferences during the contraceptive decision-making process. As such, we opted to not include this simplified system for categorizing outcomes.

3. Bleeding criteria, descriptors, and definitions

The recommended criteria include pattern, flow, and duration. The descriptors within each criterion depend on whether the contraceptive is designed to result in a predictable or unpredictable bleeding episodes. The expected pattern is a key to understanding bleeding outcomes and not whether the contraceptive is considered a short- or long-term method. The overall recommended descriptors for bleeding outcome criteria are reviewed for products designed to result in predictable episodes in Table 2 and unpredictable episodes in Table 3.

Products designed to create predictable bleeding, whether that bleeding is planned to occur monthly or less often, must be assessed in a similar manner to how a patient would evaluate the product. For example, a pill taken continuously for 11 to 12 weeks followed by a hormone-free week should have no bleeding or spotting during the 11 to 12 weeks of use and an expected bleed during the hormone-free week, because of the intention of the product design is to create predictable bleeding. However, a pill that is used daily without a hormone-free period is not expected to have predictable bleeding episodes and the evaluation methodology would be different. Product developers should be clear about the intended pattern to define bleeding outcomes appropriately and to ensure that clinicians and patients understand the intended pattern. Bleeding pattern evaluations of future contraceptive technologies should be assessed in this same manner, i.e., based on whether they are expected to have predictable or unpredictable bleeding episodes.

Flow or volume of bleeding is a separate criterion from pattern, as a person with any pattern can have no (absence of bleeding/spotting), light, moderate, or heavy flow. Pattern and flow have historically been described as separate entities for methods with expected predictable bleeding episodes.

We must also define a bleeding duration that would be considered prolonged, whether the contraceptive is intended to have predictable or unpredictable bleeding episodes. For products with a predictable pattern, this duration will also define if an expected withdrawal bleed is considered prolonged. The FIGO classification system for causes of abnormal uterine bleeding defined a normal duration of bleeding as 8 days or less based on historical studies that included hundreds of thousands of cycles throughout the entire reproductive lifespan to define normal menstrual length [22–25]. The panel that created this system agreed that the duration could still be considered abnormal for a shorter duration if the individual considered the change unacceptable [25]. We also considered a more recent study that evaluated 786 cycles in 130 participants 40 years and younger who reported regular menstrual cycles and were enrolling in barrier method contraceptive studies [26]. The investigators reported an average length of menses of 5.2 ± 1.0 days; at 2 standard deviations, 95% of cycles will be 7 days duration or less. Accordingly, we considered both 7 and 8 days for our definition of prolonged bleeding and decided that 7 days would be more appropriate for defining contraceptive bleeding outcomes.

All contraceptive products, whether they are designed to have a predictable or unpredictable bleeding pattern, will result in bleeding episodes, which are 1 or more consecutive days with bleeding or spotting. The Mishell criteria [14] describe an episode as a bleeding or spotting event that is bordered by 2 full days without any bleeding or spotting, which we agree is a clear definition.

3.1. Patterns

The bleeding pattern relates to when bleeding episodes occur. Currently available contraceptives expected to have predictable bleeding episodes include cyclic oral, transdermal, and transvaginal methods, combined injectable, and nonhormonal IUDs. Contraceptives that most commonly result in unpredictable bleeding episodes are progestin implants, hormonal IUDs, progestin-only injectable, and noncyclic progestin-only pills, patches, and rings.

Table 2
Recommended bleeding outcomes criteria and descriptors for contraceptive products with expected predictable bleeding episodes

Criteria	Descriptors	Definitions[a]
Pattern	Scheduled (withdrawal) bleeding	Bleeding/spotting episode starting during the expected interval and lasting no more than 7 d
	Unscheduled bleeding	Bleeding/spotting episode occurring at any time prior to the expected interval and not during any 7-d period defined as the withdrawal bleed
	Absence of bleeding/spotting	No bleeding or spotting during the entire cycle
Flow	None	No bleeding or spotting during the entire cycle
	Spotting	Light flow with no menstrual product use
	Lighter Usual Heavier	Flow with menstrual product use; subjective assessment as compared to typical flow when not using contraceptives
Duration[b]	Prolonged	Bleeding/spotting episode lasting more than 7 d
	Not prolonged	Bleeding/spotting episode lasting no more than 7 d
	Total number of days	Total number of bleeding, spotting, and bleeding/spotting days per cycle

[a] Will be based on the expected interval (e.g., 28-day cycle or 84-day cycle based on dosing).
[b] Duration not used when absence of bleeding/spotting occurs during the expected interval.

Table 3
Recommended bleeding outcomes criteria and descriptors for contraceptive products with expected unpredictable bleeding episodes

Criteria	Descriptors	Definitions[a]
Pattern	Absence of bleeding/spotting	No bleeding or spotting
	Infrequent	Two or less bleeding/ spotting episodes
	Normal frequency	Three or 4 bleeding/spotting episodes
	Frequent	More than 4 bleeding/spotting episodes
Flow	None	No bleeding or spotting during the entire reference period
	Spotting	Light flow with no menstrual product use
	Lighter Usual Heavier	Flow with menstrual product use; subjective assessment as compared to typical flow when not using contraceptives
Duration[b]	Prolonged	Bleeding/spotting episode lasting more than 7 d
	Not prolonged	Bleeding/spotting episode lasting no more than 7 d
	Total number of days	Total number of bleeding, spotting, and bleeding/spotting days per reference period

[a] Over a 90-day reference period.
[b] Duration not used when absence of bleeding/spotting occurs.

3.1.1. Products with an expected predictable bleeding pattern

The reference period for evaluation of products with an expected predictable pattern is based on the product design. Currently available products are commonly designed to cause 28-day or 84-day patterns, although future products could be designed with other intended cycle durations.

Those using products with expected predictable patterns may experience scheduled and unscheduled bleeding episodes and, possibly, cycles without any bleeding or spotting. Expected bleeding should occur as determined by the method and cycle duration and described as detailed in Table 2.

- Cyclic pills, patches, rings and combined injectable: Scheduled bleeding or spotting should start during the planned nonhormone use period and last no more than 7 days. Some studies have used tactics (online appendix Figure) to minimize how unscheduled bleeding is reported by including bleeding or spotting that started any time prior to the scheduled interval and continued into the interval as expected bleeding ("early withdrawal bleeding"), including some recent drug approval studies [27,28]. Another tactic is defining bleeding or spotting that starts during the hormone-free interval but continues afterward for more than 7 days ("continued withdrawal bleeding") as expected bleeding. Neither early nor continued withdrawal bleeding should be used in any bleeding descriptions or analysis plans as they do not meet expected bleeding for users, who would presume to bleed would start during the scheduled interval and not last more than a reasonable (7-day) duration. The proposed definition allows bleeding that starts during the expected interval and lasts no more than a total of 7 days to be considered as expected bleeding, thereby limiting the timing and duration of scheduled bleeding (Fig. 1).
- Nonhormonal IUDs: Scheduled bleeding is expected to occur monthly in those who experience monthly cycles and without a predictable pattern for those with previous irregular cycles. Most studies will only enroll participants with regular cycles so bleeding or spotting episodes should be reported as scheduled (monthly) and should not last for more than 7 days. All bleeding or spotting not starting during the expected interval would be reported as unscheduled bleeding. If a study includes participants with baseline unpredictable (irregular) cycles, then bleeding descriptions should be stratified to describe patterns experienced among those with baseline regular cycles and those without baseline regular cycles, with the latter being reported according to recommendations for products with an unpredictable pattern.

3.1.2. Products with an expected unpredictable bleeding pattern

Products with an expected unpredictable bleeding pattern do not have scheduled bleeding episodes so the number of days of bleeding or spotting during a reference period should be used to standardize bleeding pattern description. We recognize that no ideal reference period duration exists; examples include the historical 90-day period defined by Belsey [10], 84 days (3, 28-day cycles), and 91 days (1 year [365 days] divided into approximately 4 equal time periods). Since these time frames are relatively similar, we see no benefit from changing from current recommendations to use a 90-day period, as maintaining the 90-day period will allow comparisons across studies. As such, we recommend

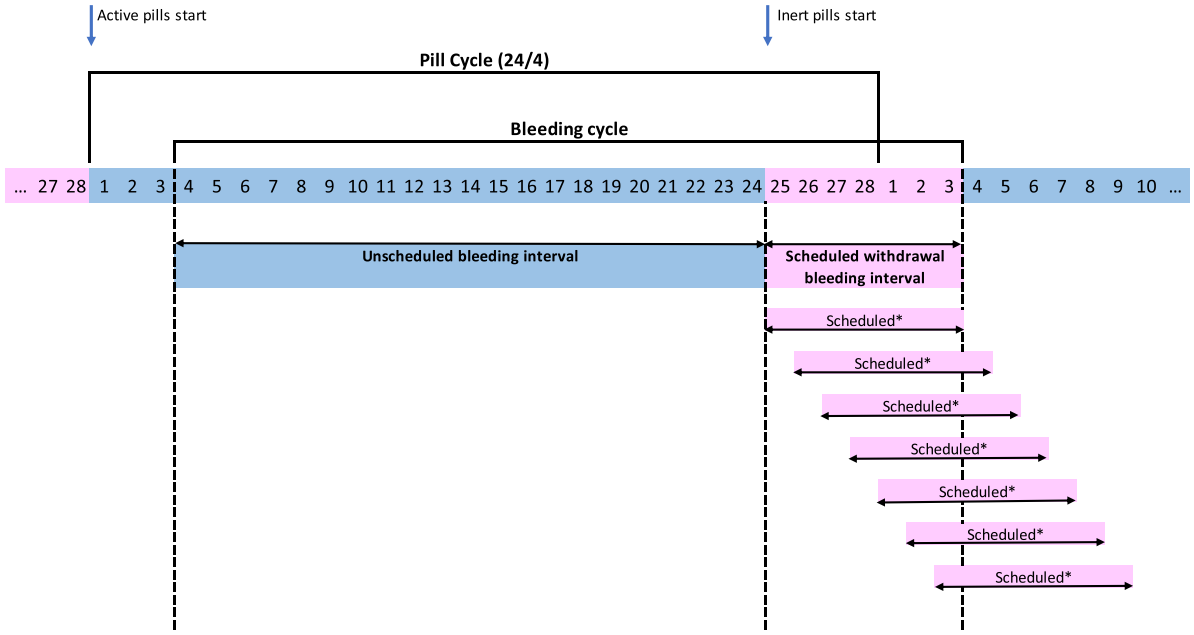


Fig. 1. Bleeding definitions for expected predictable bleeding pattern product using a 24-day hormone/4-d placebo pill example.* Scheduled bleeding can include any bleeding or spotting that starts during the scheduled bleeding interval and lasts no more than 7 days.

that contraceptive bleeding descriptions with products that are expected to have an unpredictable bleeding pattern be described in 90-day reference periods. However, we also recognize that patients do not typically think about their bleeding in 90-day intervals. Accordingly, bleeding patterns should also be described in monthly (30-day) episodes to aid with patient counseling and understanding. For example, a product that typically results in 15 days of bleeding at unpredictable intervals over a 90-day reference period should be analyzed to identify ranges of bleeding during 30-day intervals.

Pattern description for products with an expected unpredictable pattern relates to the frequency of bleeding, defined as absence of bleeding or spotting, infrequent, normal frequency, and frequent (Table 3).

3.2. Flow

Describing objective descriptors for flow is difficult because each person views their own flow subjectively, and this influences personal decisions about when to use any menstrual products as well as when and how often they change these products. Thus, any descriptors need to account for individual real-life input when assessing flow.

Because flow is independent of pattern, similar definitions for flow can be used for methods with predictable or unpredictable bleeding patterns.

3.2.1. Absence of bleeding/spotting

We use the phrase “absence of bleeding/spotting” rather than amenorrhea because the latter is a defined term related to a menstrual abnormality (3 menstrual cycles or 90 days pass without any menstrual bleeding) and not contraception-induced bleeding changes. Some contraceptives are designed with the potential to decrease or eliminate bleeding as a desired effect and, thus, these are not creating an abnormal menstrual cycle. Additionally, a single cycle with no bleeding would not meet amenorrhea criteria. As such, absence of bleeding/spotting is the preferred term to describe no bleeding or spotting over a specified time with contraceptive use.

For methods with an expected predictable pattern, absence of bleeding/spotting applies to the interval during which bleeding could occur. For example, persons using methods formulated with a 28-day cycle would be described as having absence of bleeding/spotting if no bleeding or spotting occurred for the entire 28-day cycle (Table 2). Lack of an expected withdrawal bleed (scheduled bleeding) is not absence of bleeding/spotting if bleeding occurred at other times during the cycle and should be defined clearly as lack of withdrawal bleeding. For those methods with an expected unpredictable pattern, absence of bleeding/spotting is no bleeding or spotting for the 90-day reference period, as described in Table 3.

3.2.2. Describing flow outcomes

Studies can approach defining flow amount subjectively or objectively. Objective definitions commonly define spotting and then, by default, bleeding is anything heavier than spotting. Still, spotting definitions can vary between studies, with some defining spotting as the use of no menstrual products and others using a definition of a panty liner or less. Flow definitions become more complicated with the understanding that modern menstrual products now include cups and discs and not just liners, tampons, and pads.

Defining spotting has been a challenge for decades [12]. However defined, some have recommended considering spotting as a type of bleeding (flow) and others as nonflow, suggesting that patients consider spotting as having little nuisance value [12]. It is likely that patients will vary in whether they consider spotting troublesome. Cultural, religious, and other personal factors may influence both the level of nuisance of spotting or any flow, as well as whether a patient chooses to use menstrual products.

Because regulatory agencies commonly request differentiation of spotting and bleeding, we have considered the continued importance of providing a clear definition of spotting to ensure consistency across studies. Mishell and colleagues [14] defined spotting as 'no use of menstrual products' and we agree that we should continue to use that definition uniformly. As such, we consider spotting as flow for which the person uses no menstrual products and any greater amount is bleeding. In some populations, the availability of menstrual products may be limited and thus reduce their use. When a study is performed in a population with such restrictions, the authors should describe this issue and how they choose to define spotting and bleeding for their particular population. We also strongly encourage evaluations of how people in different populations differentiate spotting and light bleeding. We expect that such descriptions will show tremendous subjective variation.

To describe bleeding flow, researchers have commonly used subjective descriptors of "light," "medium," or "heavy." Because each person can only make comparisons to their own typical flow, descriptions of flow should occur in a manner that clearly uses individual interpretation. Thus, for bleeding, flow should be described as "lighter," "usual," or "heavier" as compared to the individual's typical flow when not using contraception rather than asking individuals to describe the amount as "light," "medium," or "heavy." On days in which a study participant notes different flow amounts, for example, both spotting and lighter bleeding, the heavier flow should be included as the outcome. Should a study need to assess actual amounts, descriptors are insufficient and that study should use objective measuring tools, such as a pictorial blood loss assessment chart (known as PBAC) or alkaline hematin blood loss measurements [29,30].

3.3. Duration

Duration of bleeding for products with an expected predictable bleeding pattern refers to whether a scheduled (withdrawal bleed) or unscheduled bleeding episode is prolonged. In accordance with our earlier statement that an expected withdrawal bleed should not be longer than 7 days duration, we consider any bleeding and/or spotting episode of more than 7 days to be prolonged.

For products with an expected unpredictable bleeding pattern, the same definition for a prolonged bleeding episode applies. However, such products should also include a description of the total number of days of bleeding or spotting during each 90-day reference period.

4. Measurements

4.1. Changes over time

Many potential users of long-acting reversible contraceptives are willing to endure several initial months of unfavorable bleeding if patterns improve over time. To provide this assurance, large, long-term studies are needed that recruit subjects who represent a cross-section of society with varying demographic features. Drug approval is an expensive business with the pharmaceutical industry spending on average $1.3 billion to launch any new drug [31]. Long-term studies run the risk of participants being lost to follow-up or discontinuing before the end of the trial period because they desire pregnancy or no longer require contraception.

We suggest that bleeding analyses in users of combined hormonal contraceptive methods should include diary-based outcomes for at least 12 months whether the regimen is 21/7, 24/4, extended cycle or continuous. Studies of progestin-only injectables, given every 3 months, should ideally provide diary-based bleeding outcome data for the first 24 months as evidence suggests that prolonged bleeding and spotting decrease over time and episodes of no bleeding (defined as absence of bleeding/spotting for 90 d) increases [32,33]. Long-acting contraceptive methods should have some description of bleeding outcomes for the full duration of use so that patients can understand what to expect over time. This information is important for both clinicians and patients as some methods result in higher rates of absence of bleeding/spotting with continued use, so providing information on bleeding characteristics beyond the first year is important to document such change. While reviewing available data on current methods (online appendix Table), we recognized that commonly used absence of bleeding/spotting rates for injectables, which have an unpredictable pattern, are not reported in package labels using 90-day reference periods, as had even been recommended from the Belsey criteria. These data further underscore the need for researchers and regulatory agencies to use a single set of recommendations for reporting bleeding outcomes.

Importantly, discontinuation rates for bleeding complaints should be clearly presented for each year of the duration of use so clinicians and patients can discriminate if bleeding outcomes may be changing/improving due to study exit of those with problems.

4.2. Measuring bleeding outcomes in contraceptive trials

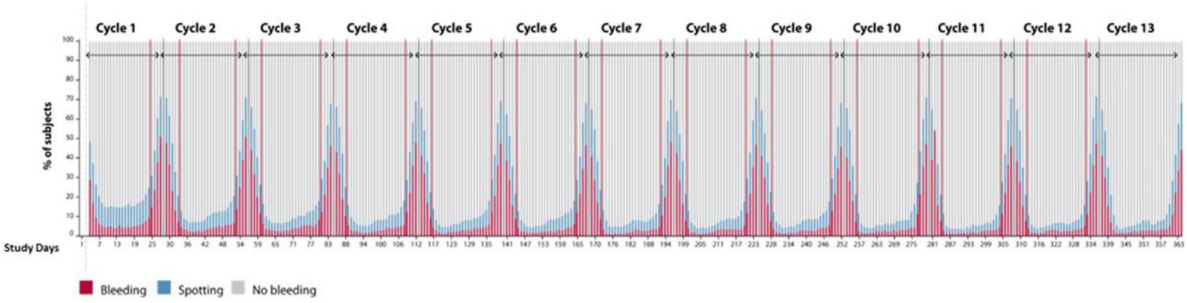
Different methodologies exist for assessment of contraceptive-induced bleeding changes in contraceptive trials. Daily diaries provide valuable information and are also used during efficacy trials to document other contraceptive use and when needed, sexual intercourse. Ideally, diaries are completed every day to summarize the experiences during that or the prior day. In all fairness, these are human trials and the volunteers should not be expected to perform in a manner that is beyond the norm. Why would we consider systems that require perfect completion of a diary daily to be both normal and reflective of a generalizable patient pool?

A big question for diaries in studies is the use of paper vs electronic diaries. Electronic diaries offer the advantage of instant participant data entry and less work to report diary information by study staff. However, electronic diary systems typically include warnings to study staff whenever a participant does not complete the diary as expected. Of course, that means study staff receive multiple notifications and then need to follow-up repeatedly with participants when diaries are not completed on schedule. Electronic diaries typically lockout participants after 1 or 2 days so data cannot be entered from a prior day, resulting in missing information.

For electronic diaries, some systems work directly via text on a participant's own smartphone. Systems that do not allow participants to use their own phones have the potential to create more lost data as participants are asked to carry 2 phones (their own and 1 for the study) with them daily. Incorporating electronic data capture into an app on a participant's own phone is ideal for this methodology.

Paper diaries offer the opportunity for study participants to fill in this information on a daily basis with the reality that some data may not be filled in every day. For a person using a method with infrequent bleeding, does that really matter? If a person has no bleeding or spotting for an entire month, does she need to report that absence of bleeding/spotting daily via an electronic system? A paper diary also allows recall (to the best of a participant's ability) to fill in missing data which an electronic system disallows. Note

a. Daily bleeding and spotting incidence over 13 cycles for a single method



b. Bleeding and spotting reported together to compare methods

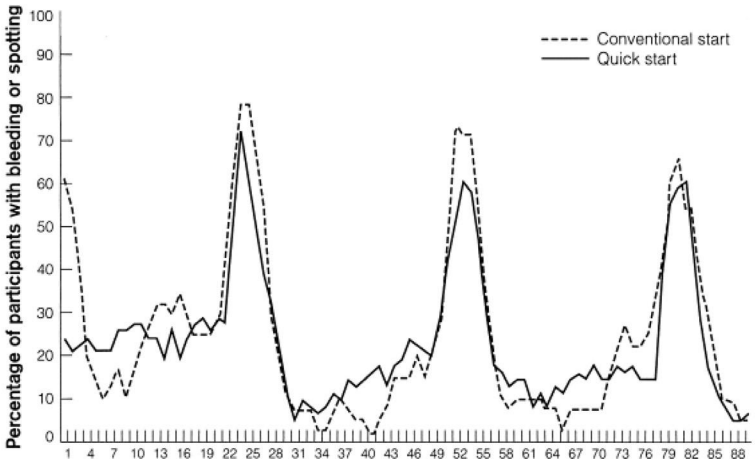


Fig. 2. Examples of graphical presentation of scheduled and unscheduled bleeding for contraceptive products with expected predictable bleeding patterns. (a) Daily bleeding and spotting incidence over 13 cycles for a single method (b). Bleeding and spotting reported together to compare methods.
[a] Reprinted (open access) from Creinin MD, Westhoff CL, Bouchard C, Chen MJ, Jensen JT, Kaunitz AM, et al. Estetrol-drospirenone combination oral contraceptive: North American phase 3 efficacy and safety results. Contraception 2021;104:222-8. doi: 10.1016/j.contraception.2021.05.002.
[b] Reprinted with permission from Westhoff C, Morroni C, Kerns J, Murphy PA. Bleeding patterns after immediate vs. conventional oral contraceptive initiation: a randomized, controlled trial. Fertil Steril 2003;79:322-9. doi: 10.1016/s0015-0282(02)04680-0.

that recall is used in all aspects of data collection, such as past medical history. Thus, providing recall is potentially better than no data at all.

Electronic bleeding diaries are preferred over paper diaries in studies that recruit young participants with heavy menstrual bleeding [34]. However, participants have demonstrated comparable data integrity with recall and prospective diaries for oral contraceptive adherence [35]. While prospective data collection via diaries may improve accuracy, the added expense and burden on study participants may not be necessary. Still, the use of retrospective recall may not be appropriate for all study populations.

In general, one method is not better than the other. There are proponents of each of these methods; the important takeaway is to realize the strengths and limitations of each and that different methods for data collection may be better or worse for different types of contraceptive products. Also important is the realization that regional variation may play a role in correct use of a method; paper may work better in one country and electronic in another country. Agencies should recognize that data collection works with either method, and investigators and participants should have a voice in which method is preferable for a particular study.

5. Reporting contraceptive-induced bleeding outcomes

Researchers should report primary outcomes in tabular form using appropriate statistics. Mean ± standard deviation should only be used to describe outcomes related to the duration of bleeding or spotting when the data is normally distributed. Otherwise, researchers should report outcomes using median and interquartile ranges. For methods with an expected predictable pattern, a graphical presentation of daily bleeding or spotting would be ideal to demonstrate unscheduled and scheduled bleeding over time (Fig. 2).

6. Conclusions

We present a basic framework that future contraceptive research trials can use to clearly describe contraceptive-induced bleeding. When available, data from completed trials could be reanalyzed using these criteria. Contraceptive-induced bleeding outcomes should be described with 3 criteria: pattern, flow, and duration. Outcome descriptors will vary based on whether the expected bleeding pattern is predictable or unpredictable. For long-acting methods, bleeding outcomes should be determined, to some extent, for the full intended duration of use.

Standardizing terminology for assessment of bleeding with contraceptive use can provide clinicians and patient's consistent information to use during counseling and decision-making. Currently, comparing across studies is difficult because of different definitions, but is also limited by different populations; the new criteria can create uniform definitions, but we will always have the limits of different populations, especially when comparing data from older trials to new ones. We encourage continued work to validate our recommended measures and provide updates, as needed, which will improve our ability to best communicate bleeding changes.

Describing pattern (which includes frequency), flow and duration as independent criteria allows a fuller understanding of contraceptive-induced bleeding changes and can allow better assessments of acceptability and continuation. As an example, a patient with infrequent prolonged bleeding may have different product acceptability than one with frequent prolonged bleeding. We do not know, because of the descriptors used to date, whether duration or frequency matters more from a descriptor standpoint for patient acceptability. These recommendations do not address whether a patient considers bleeding as bothersome/not bothersome or other measures of satisfaction/acceptability.

Contraceptive users often have clear preferences related to expectations for desired bleeding outcomes with their method [36]. Accordingly, we have tried to approach these recommendations from the patient's view, keeping in mind that patients likely expect bleeding outcomes to be equal to or better than their current patterns. Researchers can build on these recommendations by assessing patients' preferences for different contraceptive bleeding characteristics across various populations and cultures as well as the effect of bleeding on quality of life. These data can further inform counseling information and support any future recommendations.

7. Disclosure and Funding

Conflicts of interest: Mitchell D. Creinin: has received speaking honorarium from Gedeon Richter, serves on an Advisory Board for Fuji Pharma and Merck, and is a consultant for Estetra, FHI360, Libbs, Mayne, Medicines360, and Merck. The Department of Obstetrics and Gynecology, University of California, Davis, receives contraceptive research funding for Dr. Creinin from Chemo Research SL, Evofem, HRA Pharma, Medicines360, Merck, and Sebela. Carolina Sales. Vieira has received financial support to attend pharmaceutical advisory boards, speak at educational meetings/webinars/conferences, undertake scientific research from Bayer, MSD/Organon, and Exeltis. Carolyn L. Westhoff has received speaking honoraria from Mayne, serves on an Advisory Board for Mithra and Therapeutics MD, and is a consultant for HRA Pharma and for Bayer and Merck (as a DSMB member). Columbia University receives contraceptive research funding for Dr. Westhoff from Chemo Research SL, Evofem, Medicines360, and Sebela. Diana Mansour has received financial support to attend pharmaceutical advisory boards, speak at educational meetings/webinars/conferences, and undertake scientific research from Astellas, Bayer, Consilient Health, Gedeon Richter, HRA Pharma, MSD, Mithra, Organon, Pfizer, Viatris and Vifor Pharma.

Funding: The project was supported by a non-directed research grant from Merck and nondirected gifts from Cooper Surgical, Exeltis, Gedeon Richter, Medicines360, Mithra, Mayne, Sebela, and TherapeuticsMD. Funders had no input into any content and did not provide any review prior to publication.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi: 10.1016/j.contraception.2022.05.011.

References

[1] Rosenberg MJ, Waugh MS. Oral contraceptive discontinuation: a prospective evaluation of frequency and reasons. Am J Obstet Gynecol 1998;179(3):577–82 Pt 1. doi:10.1016/s0002-9378(98)70047-x.
[2] Kaunitz AM. Long-acting injectable contraception with depot medroxyprogesterone acetate. Am J Obstet Gynecol 1994;170(5):1543–9 Pt 2PMID: 8178904.
[3] Grunloh DS, Casner T, Secura GM, Peipert JF, Madden T. Characteristics associated with discontinuation of long-acting reversible contraception within the first 6 months of use. Obstet Gynecol 2013;122(6):1214–21. doi:10.1097/01.AOG.0000435452.86108.59.
[4] 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 26;15(1):114. doi:10.1186/s12978-018-0561-0.
[5] Cohen R, Sheeder J, Teal SB. Predictors of discontinuation of long-acting reversible contraception before 30 months of use by adolescents and young women. J Adolesc Health 2019;65(2):295–302 2019.02.020. doi:10.1016/j.jadohealth.
[6] Kaunitz AM, Bissonnette F, Monteiro I, Lukkari-Lax E, Muysers C, Jensen JT. Levonorgestrel-releasing intrauterine system or medroxyprogesterone for heavy menstrual bleeding: a randomized controlled trial. Obstet Gynecol 2010;116:625–32. doi:10.1097/AOG.0b013e3181ec622b.
[7] Mawet M, Nollevaux F, Nizet D, Nizet D, Wijzen F, Gordenne V, et al. Impact of a new levonorgestrel intrauterine system, Levosert(®), on heavy menstrual bleeding: results of a one-year randomised controlled trial. Eur J Contracept Reprod Health Care 2014;19:169–79. doi:10.3109/13625187.2014.894184.
[8] Sivin I, el Mahgoub S, McCarthy T, Mishell DR Jr, Shoupe D, Alvarez F, et al. Long-term contraception with the levonorgestrel 20 mcg/day (LNg 20) and the copper T 380Ag intrauterine devices: a five-year randomized study. Contraception 1990;42(4):361–78. doi:10.1016/0010-7824(90)90046-x.
[9] Rowe P, Farley T, Peregoudov A, Piaggio G, Boccard S, Landoulsi S, et al. Safety and efficacy in parous women of a 52-mg levonorgestrel-medicated intrauterine device: a 7-year randomized comparative study with the TCu380A. Contraception 2016;93(6):498–506. doi:10.1016/j.contraception.2016.02.024.
[10] 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. doi:10.1016/0010-7824(86)90006-5.
[11] 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(2):42–51.
[12] Belsey EM, Farley TM. The analysis of menstrual bleeding patterns: a review. Contraception 1988;38(2):129–56. doi:10.1016/0010-7824(88)90035-2.
[13] Mishell DR Jr, 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. doi:10.1016/j.contraception.2006.08.008.
[14] Mishell DR Jr, 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;75:11–15. doi:10.1016/j.contraception.2006.08.012.
[15] Baldaszti E, Wimmer-Puchinger B, Löschke K. Acceptability of the long-term contraceptive levonorgestrel-releasing intrauterine system (Mirena): a 3-year follow-up study. Contraception 2003;67:87–91. doi:10.1016/s0010-7824(02)00482-1.
[16] Bastianelli C, Farris M, Mirena Rapiti S. An Italian experience. Minerva Ginecol 2011;63:343–9.
[17] Modesto W, Bahamondes MV, Bahamondes L. A randomized clinical trial of the effect of intensive versus non-intensive counselling on discontinuation rates due to bleeding disturbances of three long-acting reversible contraceptives. Hum Reprod 2014;29:1393–9. doi:10.1093/humrep/deu089.
[18] Williamson PR, Altman DG, Blazeby JM, Clarke M, Gargon E. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative. Trials 2011;12:A70.
[19] Williamson P, Clarke M. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: Its Role in Improving Cochrane Reviews. Cochrane Database Syst Rev. 2012;(5):ED000041. doi: 10.1002/14651858.ED000041.
[20] Hallberg L, Högdahl AM, Nilsson L, Rybo G. Menstrual blood loss–a population study. Variation at different ages and attempts to define normality. Acta Obstet Gynecol Scand 1966;45:320–51. doi:10.3109/00016346609158455.
[21] Mansour D, Fraser IS, Edelman A, Vieira CS, Kaunitz AM, Korver T, et al. Can initial vaginal bleeding patterns in etonogestrel implant users predict subsequent bleeding in the first 2 years of use? Contraception 2019;100(4):264–8. doi:10.1016/j.contraception.2019.05.017.
[22] Munro MG, Critchley HO, Broder MS, Fraser IS. FIGO Working Group on Menstrual Disorders. FIGO classification system (PALM-COEIN) for causes of abnormal uterine bleeding in nongravid women of reproductive age. Int J Gynaecol Obstet 2011;113(1):3–13. doi:10.1016/j.ijgo.2010.11.011.

[23] 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(3):393–408. doi:10.1002/ijgo.12666.
[24] Mansour D, Korver T, Marintcheva-Petrova M, Fraser IS. The effects of Implanon on menstrual bleeding patterns. Eur J Contracept Reprod Health Care 2008;13:13–28 Suppl 1. doi:10.1080/13625180801959931.
[25] Fraser IS, Critchley HO, Munro MG, Broder M. Can we achieve international agreement on terminologies and definitions used to describe abnormalities of menstrual bleeding? Hum Reprod 2007;22(3):635–43. doi:10.1093/humrep/del478.
[26] Creinin MD, Keverline S, Meyn LA. How regular is regular? An analysis of menstrual cycle regularity. Contraception 2004;70(4):289–92. doi:10.1016/j.contraception.2004.04.012.
[27] Masson S, Franssen E, Hilditch JR, Powell MG. A clinical comparison of two triphasic oral contraceptives with levonorgestrel or norethindrone: a prospective, randomized, single-blind study. Contraception 1993;47(1):43–54. doi:10.1016/0010-7824(93)90108-j.
[28] Creinin MD, Westhoff CL, Bouchard C, Chen MJ, Jensen JT, Kaunitz AM, et al. Estetrol-drospirenone combination oral contraceptive: North American phase 3 efficacy and safety results. Contraception 2021;104:222–8. doi:10.1016/j.contraception.2021.05.002.
[29] Zakherah MS, Sayed GH, El-Nashar SA, Shaaban MM. Pictorial blood loss assessment chart in the evaluation of heavy menstrual bleeding: diagnostic accuracy compared to alkaline hematin. Gynecol Obstet Invest 2011;71:281–4. doi:10.1159/000320336.
[30] Magnay JL, O'Brien S, Gerlinger C, Seitz C. A systematic review of methods to measure menstrual blood loss. BMC Womens Health 2018;18:142 22. doi:10.1186/s12905-018-0627-8.
[31] Wouters OJ, McKee M, Luyten J. Estimated research and development investment needed to bring a new medicine to market. 2009-2018. JAMA 2020;323(9):844–53. doi:10.1001/jama.2020.1166.
[32] Hubacher D, Lopez L, Steiner MJ, Dorflinger L. Menstrual pattern changes from levonorgestrel subdermal implants and DMPA: systematic review and evidence-based comparisons. Contraception 2009;80(2):113–18. doi:10.1016/j.contraception.2009.02.008.
[33] Depo-Provera CI package insert: Pfizer. Depo-Provera prescribing information (10/2010).
[34] Jacobson AE, Vesely SK, Haamid F, Christian-Rancy M, O'Brien SH. Mobile application vs paper pictorial blood assessment chart to track menses in young women: a randomized cross-over design. J Pediatr Adolesc Gynecol 2018;31(2):84–8. doi:10.1016/j.jpag.2017.09.009.
[35] Huber LR, Broel EC, Mitchelides AN, Dmochowski J, Dulin M, Scholes D. Comparison of prospective daily diaries and retrospective recall to measure oral contraceptive adherence. Contraception 2013;88(4):492–7. doi:10.1016/j.contraception.2013.02.007.
[36] Nelson AL, Cohen S, Galitsky A, Hathaway M, Kappus D, Kerolous M, et al. Women's perceptions and treatment patterns related to contraception: results of a survey of US women. Contraception 2018;97(3):256–73. doi:10.1016/j.contraception.2017.09.010.