Document Type:	Statistical Analysis Plan
Official Title:	An open-label, parallel-group, randomized, multicenter study to assess the safety and efficacy of vilaprisan in subjects with uterine fibroids versus standard of care
NCT Number:	NCT03194646
Document Date:	04 Mar 2021

An open-label, parallel-group, randomized, multicenter study to assess the safety and efficacy of vilaprisan in subjects with uterine fibroids versus standard of care

Short title: **A**ssess **S**afety and Efficacy of Vilaprisan in Subjects with U**ter**ine Fibr**oid**s

Bayer study drug	BAY 1002670		
Study purpose:	Safety and efficacy		
Clinical study phase:	3	**Date:**	04 Mar 2021
Study No.:	16953	**Version:**	2.0
Author:	PPD		

This Statistical Analysis Plan is produced on a word-processing system and bears no signatures.

The approval of the Statistical Analysis Plan is documented in a separate Signature Document.

Table of Contents

Abbreviations

AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase (also known as GPT)
AST	Aspartate aminotransferase (also known as GOT)
B	Bleeding episode
BHC	Bayer HealthCare
BMD	Bone mineral density
BMI	Body mass index
CGI-I	Clinical Global Impression - Investigator
CI	Confidence interval
ClinRO	Clinician-reported outcome
cm	Centimeter
CSR	Clinical Study Report
DEXA	Dual energy X-ray absorptiometry
dL	Deciliter
eCRF	Electronic Case Report Form
eDiary	Electronic diary
eg	Exempli gratia, for example
EoT	End of treatment
EoFUP	End of follow-up
ePRO	Electronic patient-reported outcomes
FAS	Full analysis set
FUP	Follow-up
g	Gram
HMB	Heavy menstrual bleeding
HRQoL	Health-related quality of life
IB	Investigator's brochure
ie	id est, that is
INR	International normalized ratio
IVRS/IWRS	Interactive voice/web response system
MedDRA	Medical Dictionary for Regulatory Activities
METW	Moving evaluation time windows
mg	Milligram
mL	Milliliter
MP	Menstrual pictogram
N/A	Not applicable
PAEC	Progesterone receptor modulator-associated endometrial changes
PD	Pharmacodynamic(s)
PK	Pharmacokinetic(s)
PRM	Progesterone receptor modulator
PRO	Patient-reported outcome
PT	Preferred term
RND	Randomization
RAVE	Electronic data capturing system
SAE	Serious adverse event
SAF	Safety analysis set
SAP	Statistical Analysis Plan
SAS	Statistics Analysis System
SCR	Screening
SD	Standard deviation

SOC	System organ class
SoT	Start of treatment
TEAE	Treatment-emergent adverse event
TP	Treatment Period
trt	Treatment
TVU	Transvaginal ultrasound
UF-DBD	Uterine Fibroid Daily Bleeding Diary
UF-DSD	Uterine Fibroid Daily Symptom Diary
UFS-QoL	Uterine Fibroid Symptom and Quality of Life questionnaire
ULN	Upper limit of normal
US	United States
VPR	Vilaprisan
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary
wk	Week

1. Introduction

Uterine fibroids are benign tumors originating from smooth muscle cells of the myometrium. The pathophysiology of fibroids is not well understood. Genetic predisposition, exposure to steroid hormones, and growth factors play a role in formation and growth. Uterine fibroids typically appear and grow during reproductive years, but stabilize or regress after menopause. Therefore, they rarely require treatment after menopause. Clinically, uterine fibroids and associated symptoms are most prominent in the late reproductive years. The most common symptoms of uterine fibroids are heavy menstrual bleeding (HMB) and pelvic discomfort.

Uterine fibroids are the leading cause for hysterectomy. Hysterectomy is the only definitive treatment and eliminates the possibility of recurrence. Increasingly, more women desire to avoid hysterectomy, electing for a uterine preserving procedure. The surgical treatment options are numerous, and each carries both the risks for surgery itself, as well as the possibility that the woman may require subsequent surgery as new fibroids often develop over time and become symptomatic. Not surprisingly, many women would prefer not to have surgery at all. Therefore, there is a great medical need for effective pharmacological treatment suitable for long-term treatment of uterine fibroids.

The development of selective progesterone receptor modulators (PRMs) offers the potential for a novel, well tolerated medical treatment approach for women who are experiencing symptoms caused by their fibroids. Various studies have demonstrated the steroid-dependence of fibroid growth and that progesterone has a critical role.

This clinical study is part of the development program for vilaprisan. The rationale of the study is to assess the efficacy and the safety of vilaprisan in subjects with uterine fibroids in comparison to nonhormonal medical treatment in accordance with local standard of care.

The Statistical Analysis Plan (SAP) is based on following protocol versions:

- Original protocol, Version 1.0, dated 17 MAR 2017
- Amendment 03 forming integrated protocol Version 2.0, dated 13 JUN 2017
- Amendment 04 forming integrated protocol Version 3.0, dated 13 SEP 2017
- Amendment 06 forming integrated protocol Version 4.0, dated 04 JUL 2018
- Amendment 07 forming integrated protocol Version 5.0, dated 20 AUG 2018
- Amendment 09 forming global protocol Version 6.0, dated 11 DEC 2018
- Amendment 10 forming integrated protocol Version 7.0, dated 21 NOV 2019
- Amendment 11 forming integrated protocol Version 8.0, dated 17 FEB 2020
- Amendment 12 dated 06 OCT 2020 (local amendment, valid for Turkey only)

2. Study Objectives

The primary objective of this study is to evaluate the safety of vilaprisan in subjects with uterine fibroids in comparison to nonhormonal medical treatment in accordance with local standard of care.

The secondary objective of this study is to evaluate the efficacy of vilaprisan in subjects with uterine fibroids in comparison to nonhormonal medical treatment in accordance with local standard of care. With the implementation of Protocol Amendment 10 (version 7.0), additional focus will be put on safety evaluations of the endometrium, adrenal glands and skin.

Other objectives of this study are to evaluate the variability in exposure in relation to the efficacy and safety for vilaprisan and to collect patient-reported outcome (PRO) and clinician-reported outcome (ClinRO) data.

3. Study Design

This is an open-label, parallel-group, randomized, multi-center study.

The study was planned to be conducted in Europe, North America, South America, South Africa, and Asia Pacific.

Originally it was planned to randomize 1050 subjects into the study. However, recruitment into the study and treatment in the study were stopped earlier than planned.

An overview of the originally planned study design (before the modifications in study design introduced with Amendment 9, version 6.0 and subsequent amendments) is shown in Figure 3-1.

Eligible subjects were randomized to one of three vilaprisan treatment groups (Groups A1, A2, or A3) or to the standard of care group (Group B).

The planned treatment duration was about one year (Subgroup 1) or two years (Subgroup 2), respectively. The assignment to the 2 subgroups was done in a sequential manner.

After the end of treatment (EoT), subjects were planned to be followed up for 24 weeks (Subgroup 1) or for 12 weeks (Subgroup 2).

With the implementation of protocol amendment 9 (Version 6.0), no further subjects were recruited. Subjects that started to a regimen with 3-months treatment courses should finish the treatment course as planned. Subjects that started to a regimen with treatment courses of 6 months duration should finish the first 3 months, if they are currently in the first 84 days of that treatment course, finish the second 3 months, if they are currently beyond day 84 of the treatment course. Thus, with the implementation of protocol amendment 9 and the originally foreseen study design outlined in Figure 3-1 is not applicable anymore for vilaprisan groups. Subjects in treatment group B can continue the study course following the current valid version of the protocol.
With the implementation of Protocol Amendment 10 (version 7.0), this study design as outlined in Figure 3-1is no longer valid. No subjects will receive further study drug treatment. All subjects who were randomized and started treatment before the temporary pause were asked to have a comprehensive safety evaluation (with particular focus on endometrial, adrenal and skin safety) performed. This also applies to subjects who have completed or discontinued the study before or during the temporary pause, provided they have taken at least one dose of study medication.

Likewise, subjects randomized to treatment group B who are still in the study will be asked to undergo these procedures.

At the time study treatment was terminated, no subjects in subgroup 2 had reached the second year of treatment. Therefore, the statistical analysis will include patients from both subgroups without any distinction.

Figure 3-1 Design overview

Subgroup 1 (about 1 year of treatment)

A1	SCR up to 90 days	RND	12 wk trt	*B*	12 wk trt	*B*	12 wk trt	B	12 wk trt	24 wk FUP

A2	SCR up to 90 days	RND	24 wk trt	*B*	*B*	24 wk trt	24 wk FUP

A3	SCR up to 90 days	RND	12 wk trt	*B*	*B*	12 wk trt	*B*	*B*	12 wk trt	24 wk FUP

B	SCR up to 90 days	RND	Symptomatic nonhormonal medical treatment as determined by the investigators and/or watch and wait	24 wk FUP

Subgroup 2 (about 2 years of treatment)

A1	SCR up to 90 days	RND	12 wk trt	*B*	12 wk trt	*B*	12 wk trt	B	12 wk trt	*B*	12 wk trt	*B*	12 wk trt	*B*	12 wk trt	*B*	12 wk trt	12 wk FUP

A2	SCR up to 90 days	RND	24 wk trt	*B*	*B*	24 wk trt	*B*	*B*	24 wk trt	*B*	*B*	24 wk trt	12 wk FUP

A3	SCR up to 90 days	RND	12 wk trt	*B*	*B*	12 wk trt	*B*	*B*	12 wk trt	*B*	*B*	12 wk trt	*B*	*B*	12 wk trt	*B*	*B*	12 wk trt	12 wk FUP

B	SCR up to 90 days	RND	Symptomatic nonhormonal medical treatment as determined by the investigators and/or watch and wait	12 wk FUP

Groups A1, A2, and A3 treated daily with vilaprisan 2 mg during treatment periods
Group A1: 4/8 treatment periods of 12 weeks, each separated by 1 bleeding episode (3/1 regimen)
Group A2: 2/4 treatment periods of 24 weeks, separated by 2 bleeding episodes (6/2 regimen)
Group A3: 3/6 treatment periods of 12 weeks, each separated by 2 bleeding episodes (3/2 regimen)
Group B treated with standard of care (SoC) symptomatic nonhormonal medical treatment as determined by the investigators and/or watch and wait.
B = bleeding episode; FUP = follow-up; RND = randomization; SCR = screening; trt = treatment; wk = week.

The **primary safety variable** is the percentage change in bone mineral density (BMD) of lumbar spine from baseline to about one year after start of treatment (SoT) (please refer to section 6.5.1 for details). With the implementation of Protocol Amendment 10 (version 7.0), the study is being prepared for closure. A thorough safety monitoring is performed in subjects before this closure, including a recovery BMD scan. All subjects who received vilaprisan treatment will be asked to undergo a BMD scan at least 6 months after the last intake of vilaprisan, to document absence of

bone loss or adequate recovery after the end of treatment. Subjects randomized to treatment group B who are still in the study with implementation of Protocol Amendment 10 (version 7.0) should also undergo a BMD scan at the safety closeout visit, unless they had a BMD scan performed within the last 3 months prior to the safety closeout visit.

4. General Statistical Considerations

4.1 General Principles

The statistical evaluation will be performed by using the software package SAS release 9.4 or higher (SAS Institute Inc., Cary, NC, USA). All variables will be analyzed by descriptive statistical methods. The number of data available, mean, standard deviation, minimum, median and maximum will be calculated for metric data. 1st and 3rd quartiles will be provided for certain variables, which will be specified in the respective subsections. Frequency tables (counts and percentages) will be generated for categorical data. The category 'missing' will be included as a separate category in frequency tables, if applicable.

Tables, figures and listings (TLFs) will show the four treatment groups as defined in the randomization: A1, A2, A3 and B, and the corresponding total. In addition, all vilaprisan treatment groups will be displayed as Total VPR.

For time interval-based analyses of efficacy variables (except for fibroid surgeries), including analyses for the treatment break and FUP, analyses for the FUP will be based on data from subjects who have started the planned last TP. Data from FUP for subjects who did not start the planned last TP for any reason will be included in the analysis of the break period. For those subjects only efficacy-related data up to a theoretical end of the break will be considered (see 'break period' in Section 4.5 for details).

4.2 Handling of Dropouts

A subject who discontinues study participation prematurely for any reason is defined as a "dropout," if the subject has already been randomized. Drop-outs in this study will not be replaced. The details for the handling of missing data due to drop-outs are described in Section 4.3.

4.3 Handling of Missing Data

All missing or partially missing data will be presented in the subject data listing as they are recorded on the Case Report Form (CRF) and electronic diary (eDiary).

When appropriate and for a limited number of variables, certain rules described below will be implemented to avoid exclusion of subjects from statistical analyses due to missing or incomplete data.

Bleeding data

For bleeding pattern analyses, missing eDiary will be imputed based on the sponsor's standard procedures [11]. Missing bleeding data in the subject diaries will be imputed. For the uterine fibroid daily bleeding diary (UF-DBD), the worst case approach will be used for missing bleeding intensity

for the single or two consecutive missing days (ie, the maximum of the bleeding intensities of the day before and the day after the missing day[s] will be assumed).

After having imputed missing values in the UF-DBD, missing menstrual pictogram (MP) values are imputed in the next step.

For days with bleeding intensity of "mild" or higher in the UF-DBD, missing MP values will be replaced by the mean value of MP values of the days with the same bleeding intensity. In case there are no MP values with the same bleeding intensity, the mean of the MP values of next higher intensity will be used. In case there is no intensity higher than the bleeding intensity of the missing MP values available, no replacement for the missing MP values will be done for such days.

4.4 Interim Analyses and Data Monitoring

With the implementation of the protocol amendment 10 (version 7.0) a safety and efficacy analysis is planned after all subjects have completed their treatment period. All data until 10Feb2021 will be cleaned and included in the analysis. The complete primary safety data and efficacy data has been collected until this point in time and will be used for the primary safety analysis and the efficacy analysis. Therefore, the primary safety analysis and efficacy analyses will be final at this point in time. Data of the follow-up period available at that time point will also be included in the analysis.

Remaining data of the follow-up phase will be analyzed after all subjects have completed the Safety closeout visit. For all other data this safety and efficacy analysis is considered the final analysis.

Data collected in the 5-year follow-up period after the safety closeout visit will be listed and reported in an addendum to the study report. This is only applicable for Turkey.

4.5 Data Rules

Reference Start/End dates

For vilaprisan (VPR), the reference start and end dates refer to the first and last study drug intake and will be identified based on the eCRF entry (EXCATN=1). For subjects with study drug exposure unknown, the randomization date will be used as the reference start date, and the end date of last scheduled/ unscheduled visit according SV (Subject Visits dataset) will be used as the reference end date.

Missing or partially missing eCRF end dates will be replaced by the date of the last study drug intake according to the diary. In case there is no study drug intake recorded in the diary, the date of the last available visit/contact will be used, which is the maximum of the end date of last scheduled/ unscheduled visit according SV (Subject Visits dataset), and the "last visit" as collected on the eCRF page for Disposition dataset. For subjects randomized to VPR who did not take any treatment (FAS not SAF), the last available visit date will be used as the reference start date and the reference end date.

For treatment group B the reference start date is defined by the randomization date. The reference end date is defined by the start date of the End of Treatment (EoT) visit (unscheduled EoT visits not considered). Missing EoT visit dates will be replaced by the end date of the last available visit before the safety closeout visit (refer to Table 4-1). If a subject in treatment group B has no safety

closeout visit, missing EoT visit dates will be replaced by the date of last scheduled/ unscheduled visit according SV (Subject Visits dataset).

For both VPR and treatment group B, the reference end date after the missing replacement with the rules defined above should be no earlier than the reference start date. Otherwise, the reference end date will finally be replaced by the reference start date.

For any analysis, if date of first study drug exposure/ date of last study drug exposure is used, the reference start/end date for treatment group B defined as above will be used for the analysis in treatment group B, if applicable.

Table 4-1 Replacement rules for missing reference start/end dates and TP start/end dates

	Choice	VPR regimens	Treatment group B
Reference start date	1st	eCRF	Randomization date
	2nd	a. If study drug exposure unknown, use randomization date b. If no study drug exposure (FAS not SAF) use SV last visit end date	
Reference end date / Last TP end date	1st	eCRF	EoT visit start date
	2nd	Diary	Last visit before the safety closeout visit
	3rd	Max (SV last visit end date, "last visit" from eCRF)	SV last visit end date
	4th	If no study drug exposure (FAS not SAF) use SV last visit end date	
		Reference end date will finally be replaced by maximum of reference start date and reference end date.	
TP start date	1st	eCRF (TP1: reference start date) Exception 6/2 regimen: split protocol planned TPs with 168 days, assign start and end dates of these with CRF collected for planned TPs and derived start and end dates of split TPs.	TP1: reference start date TPs will be defined as consecutive 84-day time intervals starting with the date of randomization.
TP end date (excluding last TP)	1st	eCRF exception 6/2 as above	
TP end date (last TP)	1st	Reference end date	

eCRF = electronic case report form; EoT = end of treatment; SV=subject visit; TP = treatment period; VPR = vilaprisan

Treatment period

For subjects in treatment groups A1, A2, and A3, each treatment period (TP) consists of three subsequent 28-day periods. In case of premature discontinuation of study drug the last 28-day period of the respective treatment period might be shorter than 28 days. If a subject took the study drug for more than 84 days of a treatment period, the respective treatment period will be longer than 84 days, in contrast. Subjects treated in 3/1 and 3/2 regimens, have the first break after TP 1. Subjects treated in 6/2 regimens have two planned TPs before the first break.

The definition of TP here described which will be used in the analysis, differs from the definition of TP used in the protocol and Section 3 to describe the 6/2 treatment regimen. The definition used in the analysis allows for a better comparison between treatment regimens and is consistent with the reporting of other studies of the VPR project.

The first TP starts with the reference start date, the last TP ends with the reference end date. Start and end of a TP will be identified based on eCRF entries (EXCATN=1).

In 6/2 regimen, the protocol plans for TPs with 168 days each. The CRF only collects start and end dates of these protocol planned TPs.

For the analysis, the planned protocol TPs are split into 2 Analysis TPs:

- The first Analysis TP is assigned first available odd TP number and ends with day 84 (or the maximum number of available days if less than 84 days are reported).
- The second Analysis TP is assigned the next higher even TP number, starts with day 85 and ends with the last day reported in the original protocol TP, even if this exceeds a total of 168 days.

For example, protocol TP1 = analysis TP1 + analysis TP2. Protocol TP2 = analysis TP 3 + analysis TP 4 etc.

If the original planned protocol TP is less than 85 days, no such second evenly numbered Analysis TP will be assigned. As a result of this a subject in 6/2 regimen might have Analysis TP 1, TP 3 and TP 4, but no Analysis TP 2.

If the end date of the current treatment period is equal to the start date of the next treatment period based on CRF entries, it is assumed that the current treatment period ended one day earlier. The end date of the current treatment period is therefore moved back by one day.

For all TPs excluding the last TP, in case end of TP date is still missing, it will be replaced by the date of the last study drug intake within the respective TP according to diary (EXCATN=5 and EXMEDN>=1).

Subjects in treatment group B don't have any breaks. In the treatment group B TPs will be defined as consecutive 84-day time intervals starting with the date of randomization.

Replacement rules for missing end date of the last TP are analogue to those described for reference end dates.

TPs will be numbered sequentially.

Time intervals

Due to the complex visit schedule of the studies, a time-interval based approach is chosen to present data over time. Therefore, time intervals are defined which are used for different analysis topics.

The selection of time intervals used for a certain analysis topic will be described within each section.

An overview of the time intervals is given in Figure 4-1.

Figure 4-1: Time interval overview

BL/Post baseline	BL			Post baseline																					
Intervals A	BL			Treatment Phase															FUP Phase						
Intervals B	BL			TP1 + BR1						TP2 + BR2						TPx			FUP1+2					FUP ...	
Intervals C	BL			TP1			BR1			TP2			BR2			TPx			FUP1			FUP2		FUP ...	
Intervals D	BL 1-28	BL 29-56	BL 57-x	TP1 1-28	TP1 29-56	TP1 57-84	BR1 1-28	BR1 29-56	BR1 ...	TP2 1-28	TP2 29-56	TP2 57-84	BR2 1-28	BR2 29-56	BR2 ...	TPx 1-28	TPx 29-56	TPx 57-84	FUP 1-28	FUP 29-56	FUP 57-84	FUP ...	FUP ...	FUP ...	FUP ...	FUP ...
AE Intervals	Pre-treatment			"on" (TP1 + 8d)			"off"			"on" (TP2 + 8d)			"off"			"on": TPx + 8d			"off"	Post-treatment (Start: day 61 after EoT)						
Exit examination	BL, last value within interval																		Exit examination (Start: EoT -7 days); last value within this interval						

Abbreviations: BL: Baseline, TP: Treatment Period, BR: Break (not applicable for Treatment group B), FUP: Follow-up, d: days, m: months.

Treatment periods and breaks will be identified as described above. Baseline/Pre-treatment phase starts with the date of informed consent and ends on the day before the start of the Treatment phase.

The Treatment phase starts and ends with the reference start and end date. The FUP phase starts on the day after the end of the treatment phase and ends on the maximum of the last date at which a visit took place (visit end date), and the "last visit" as collected on the eCRF page. For patients who did not complete the planned schedule of TPs, a break period will be defined after the end of the last TP, as described in the previous subsection 'Break period'. This break period will be followed by the FUP.

Baseline, Treatment phase, FUP phase, TPs and breaks are subdivided into subsequent periods of different lengths (28, 84, or 168 days). Exit examination time intervals include the baseline measurement, i.e. the last non-missing value before reference start date, and the exit examination which is defined as the last non-missing value at or after the reference end date – 7 days, which is per definition the safety closeout visit for subjects who perform this visit.

For treatment group B, consecutive TPs (84-day time intervals) in the treatment phase will be used for intervals B and intervals C. Each TP will be subdivided into subsequent consecutive 28-day periods for intervals D. For intervals for assessing AEs in the treatment phase, consecutive TPs will be ended at day 60 after the reference end date. For the assessment in the follow-up phase, same rules will be used as VPR treatment groups.

Details on time intervals for AE presentation are given in Section 6.5.3.5.

Details on time intervals for BMD presentation are given in Section 6.5.1.

Break period

All VPR treatment groups include a treatment-free period following the end of TPs in 3/1 and 3/2 regiments, or end of TPs 2, 4, 6 etc. in 6/2 regimen. The length of the break period will vary depending on whether 1 or 2 bleeding episodes take place. In any case, the break period starts the day following the date of stop date of study drug administration in a particular TP, and ends the day before the study drug administration starts in the next TP.

If a subject did not start the next planned TP, due to the clinical hold or any other reason, there is no end date for the break period. In this case, a theoretical end date will be generated. For this purpose, information on the start of the previous TPs will be used. If a subject started TP1 at the i-th Day of her bleeding episode, the theoretical start date of TP2 will be assumed as the i-th Day of the bleeding episode which would have triggered re-starting medication during the treatment break (first or second bleeding depending on treatment regimen). The theoretical end date of the break period will then be assumed as the day before the theoretical start date of TP2. For TPs beyond TP2 the average of the i-th Days of the menstrual cycles when previous TPs started will be calculated and considered as start of the theoretical absent TP. This definition for theoretical end date for the break period depends on the subject experiencing a bleeding episode following the end of the TP. If no bleeding episode starts (spontaneous or induced), then the break period is considered until the end of subject observation.

This theoretical end date will be used in the analysis of number of bleeding days.

For subjects in the 6/2 regimen the break period will follow the end of TP2 or any TP with an even number. If it is necessary to define a theoretical end of the break period, as described in the previous paragraph, this will only occur if the subject started TP2 or any TP with an even number. But for the calculation of the variable number of bleeding days, if a subject withdrew while receiving treatment,

in order to use all information available at the time of analysis, the theoretical end date for the break period of the incomplete TP will be generated regardless the TP with an even or odd number.

Baseline for general parameters

In general, the last valid, non-missing value prior to the Reference start date will be considered as baseline value. For variables analyzed by 28-day periods, the baseline period corresponds to the last 28 days before the reference start date. This definition does not apply to analyses related to menstrual blood loss where a specific definition for baseline menstrual blood is defined (see baseline menstrual blood loss for 28 days).

Bleeding caused by conducting endometrial biopsy

If an endometrial biopsy was conducted, bleeding on the day of intervention and the 3 days thereafter will not be considered in the bleeding related evaluation, i.e., bleeding/spotting on the day of biopsy and 3 days thereafter will not be considered as day(s) with bleeding/spotting and MP values will be set to 0 mL. This rule is applicable for all bleeding related efficacy analyses.

A biopsy will be regarded as conducted if the intervention was attempted, either successfully or unsuccessfully. Successful means that the date of biopsy is given and 'Biopsy sample obtained - Not done' is not ticked in the eCRF. Unsuccessful means that the date of biopsy is given, 'Biopsy sample obtained - Not done' is ticked in the eCRF, and the reason 'Unsuccessful attempt' is selected.

Baseline menstrual blood loss for 28 days

Baseline menstrual blood loss for 28 days is defined as the sum of volume of menstrual blood loss data of respective days in the baseline cycle. If baseline cycle length is more than 28 days, only data from the first day of cycle to the 28th day of cycle will be used. If baseline cycle length is less or equal than 28 days, data from the first day of cycle to the last day of cycle will be used.

Baseline menstrual cycle

The baseline menstrual cycle length is defined as the days from the first day of cycle until the last day of cycle. First day of the baseline cycle is Day 1 of the bleeding episode following screening Visit 1. Last day of cycle is the day preceding the next bleeding episode. This definition will apply, if during this cycle the inclusion criterion of HMB diagnosis is fulfilled (menstrual blood loss > 80 mL). If HMB is not confirmed in the first bleeding episode following screening Visit 1, then the second bleeding episode will be considered. Similarly, the first day of the second cycle is the Day 1 of the second bleeding episode following screening Visit 1 and the day preceding the next bleeding episode will be considered the last day of the cycle. Again, this second bleeding episode will be considered the baseline cycle only if the HMB criterion is fulfilled, otherwise the next bleeding episode, if recorded, will be inspected for the diagnosis of HMB. If none of the bleeding episodes preceding the first drug administration/ reference start date fulfill the HMB inclusion criterion, the first bleeding episode following screening Visit 1 will be considered for the baseline cycle. If no such bleeding episode exists, the bleeding episode already ongoing at screening Visit 1 will be considered for the baseline cycle, starting with Day 1.

During the screening period the diagnosis of HMB will be inspected using the definition of bleeding episode based on the MP, as explained in this subsection and in the protocol. In contrast, during the analysis only the definition of bleeding episode based on the UF-DBD will be used. Therefore, the

baseline cycle defined during the analysis here described is independently derived and may differ from the bleeding episode used to assess the HMB diagnosis during the screening period.

Bleeding episode (assessed by the UF-DBD)

Bleeding episode is defined as day(s) with bleeding/spotting of which at least one day is of intensity "mild" or higher, preceded and followed by at least 2 days with diary entry: "no vaginal bleeding" (in case the first bleeding episode starts directly after Visit 1, the preceding 2 bleed-free days may not be recorded for this first bleeding episode).

This definition of bleeding episode, based on the UF-DBD, is the only bleeding definition implemented in the analysis to identify bleeding episodes. The protocol includes a definition of bleeding episode based on the MP, which will be implemented during the screening period to facilitate the assessment of the HMB inclusion criterion (criterion 4, MBL >80 mL) by clinical sites. This latter definition of bleeding period will not be used in the analysis.

Menstrual blood loss for 28 days

Menstrual blood loss for 28 days is defined as the sum of menstrual blood loss [mL] by MP in a 28-day period. Non-overlapping 28-day periods will be created along the different study periods: TPs, break periods (if applicable), and FUP phase. In case of premature discontinuation of study drug the last 28-day period of the respective treatment period might be shorter than 28 days. For details about the volume of MBL assigned to MP, see Appendix 4.

Volume of menstrual blood loss per bleeding episode after the reference end date

Volume of menstrual blood loss for a bleeding episode is the sum of volume of menstrual blood loss [mL] data by MP in the respective episode after the reference end date.

Amenorrhea (defined as < 2 mL during last 28 days of treatment course) assessed by MP method

Amenorrhea is defined as menstrual blood loss less than 2 mL during the last 28 days of the respective treatment period. For missing data replacement please refer to section 4.3.

i) the sum of the non-missing (i.e. original or imputed) MP values is ≥ 2 mL, or
ii) there is at least one day with missing MP value and bleeding intensity of "mild" up to "very severe" after having applied the imputation rules, or
iii) the subject did not complete at least 8 weeks of treatment.

If neither of i), ii) or iii) applies, the subject will be considered as having amenorrhea.

The last days under treatment with 28 days with non-missing (i.e. original or imputed) MP data will be used to calculate whether amenorrhea occurred. Thus, if the last 28 days of treatment include days with missing MP data, the left margin of this interval will be displaced backwards in time and step-wise until the required number of 28 non-missing values are included.

The left margin may be displaced backwards at maximum up to the end of the first bleeding episode when treatment was started in respective treatment period. For TP with an even number in the 6/2 regiment, the left margin may be displaced backwards at maximum up to the end of the first bleeding episode of the preceding odd-numbered TP from the same protocol defined 168-day TP when treatment was started.

Absence of bleeding (spotting allowed) during the last 28 days of the treatment; based on the UF-DBD

Absence of bleeding is defined as no bleeding (spotting allowed) during the last 28 days of the respective treatment period based on the UF-DBD. For missing data replacement please refer to section 4.3.

The last days under treatment with 28 days with non-missing bleeding eDiary data will be used to calculate whether absence of bleeding occurred. Thus, if the last 28 days of treatment include days with missing bleeding eDiary data, the left margin of this interval will be displaced backwards in time and step-wise until the required number of 28 non-missing values are included.
The left margin may be displaced backwards at maximum up to the end of the first bleeding episode when treatment was started in respective treatment period. For TP with an even number in the 6/2 regiment, the left margin may be displaced backwards at maximum up to the end of the first bleeding episode of the preceding odd-numbered TP from the same protocol defined 168-day TP when treatment was started.

Subjects who discontinue treatment prior to completion of 8 weeks of treatment in the respective TP will be considered as not experiencing absence of bleeding.

Number of bleeding days

Number of bleeding days is defined as number of bleeding days (based on the UF-DBD) from Day 1 of the first treatment period until one day before the theoretical next treatment period would start after the last study treatment period. Definition of the day before theoretical next treatment period end date, i.e., theoretical end date for the break period, refers to the rules defined in break period above.

All information available at the time of analysis will be included. For example, for a subject allocated to treatment group A1 (3/1 regimen) in subgroup 1 who completed the full planned treatment, the number of bleeding days will be counted from the 1st day of TP1 until the date when a new TP would theoretically start following the end of TP4 and the required bleeding episode that follows TP4, without gaps. If a subject withdrew while receiving treatment in TP3, the number of bleeding days will be considered from the 1st day of TP1 until the date when a new TP would theoretically start following the end of TP3 and the required bleeding episode that follows TP3, irrespective of the length of the incomplete TP3.

For subjects who did not start the next planned TP, due to the study being temporarily paused or any other reasons, as described in the section ‘Break period’, a theoretical end of break period will be derived in these cases. The number of bleeding days will be counted until the derived end of break period.

The period for evaluation of “number of bleeding days” for the treatment group B starts on the reference start date and ends on the reference end date. A theoretical end of break period following the reference end date will not be generated.

The analysis of the number of bleeding days from these pre-determined evaluation periods will be normalized to 28 days and 365 days.

Onset of amenorrhea for analysing time-to-event data

Onset of amenorrhea is defined by the first day for which the menstrual blood loss (assessed by MP) for all subsequent 28-day periods up to the end of the treatment period is <2 mL. Amenorrhea status will be assessed for each of the subsequent, overlapping 28-day periods within a treatment period. The censoring mechanism is assumed to be non-informative and the subjects will be handled as right-censored, if applicable. If a subject completed the 84 days of a treatment period and did not experience an onset of amenorrhea, the subject will be censored on Day 57 of the treatment period.

For prematurely discontinued subjects, censoring also depends on whether an onset of amenorrhea can be ruled out or not during the time the subject was in the respective treatment period. An onset of amenorrhea can be ruled out, if in the last 28-day period before the subject discontinued the respective treatment period, the sum of MP values exceeds 2 mL. If this is not the case, then an onset of amenorrhea cannot be ruled out for the respective treatment period. Thus, there are in principle the two possibilities:

1. In case the subject discontinued the respective treatment period prematurely and an onset can be ruled out, subjects will be treated as right-censored at the first day of the last 28-day period after which MP bleeding values for all further days are missing.

2. If the subject prematurely discontinued the respective treatment period but an onset cannot be ruled out, the subject will be censored on the day before the first day of the first 28-day period which indicates that an onset may have taken place.

Onset of controlled bleeding

Defined by the first day for which the menstrual blood loss (assessed by MP) for all subsequent 28-day periods up to the end of the treatment period is <80 mL. The first 28 days with non-missing MP values until the last 28 days with non-missing MP values under treatment in the respective treatment period will be considered for deriving the time to onset of controlled bleeding. Time to onset of controlled bleeding is derived in the same way as time to onset of amenorrhea. The only difference is the use of a less strict criterion (i.e. <80 mL instead of <2 mL).

Start of bleeding after last study drug intake (assessed by UF-DBD)

Defined by the first day for which bleeding intensity is “mild” or higher after the last study drug intake. If no start of bleeding after last study drug intake is documented,

- Subjects will be censored at latest available date from the UF-DBD.
- If bleeding has to be induced, the subject´s data will be censored at the time point of bleeding induction. The bleeding induction information will be obtained from the concomitant medication page.
- For the analyses after last study drug intake in TPx, the subject will be censored at latest at TPx+1 start date if she starts TPx+1, where x=1,…, n.

Largest fibroid

The largest fibroid is defined as the fibroid which is the largest at baseline (measured by ultrasound). The fibroid identified at baseline as the largest will be followed up at each scheduled visit.

Fibroids volume measured by ultrasound (mL) [(2)] $\frac{\pi \times a \times b \times c}{6}$);

a = Largest diameter of the fibroid (cm), b = First diameter perpendicular to largest diameter (cm), c = Second diameter perpendicular to largest diameter (cm)

If only 2 diameters are available, the formula will be the same but the third diameter will be replaced by the mean of the available two diameters.

If only 1 diameter is available, the formula will be: $= \frac{4\times\pi\times r^3}{3}$; $r = \frac{\text{available diameter of the fibroid (cm)}}{2}$

If one or two diameters are recorded as 0 cm, then those diameters will be considered missing and the rules above will apply.

Volume of 3 largest fibroids measured by ultrasound (mL)

Defined by the sum of the 3 largest fibroid volumes. If less than 3 fibroids are present, the volume of the actual number of fibroids will be summed. The 3 fibroids identified at baseline as the largest will be followed up at each scheduled visit.

Uterine volume measured by ultrasound(mL) $= \frac{\pi\times a\times b\times c}{6}$;

a = Maximum width anteroposterior of uterus (cm), b = Maximum width transverse (cm)of uterus, c = Corpus + cervix length (cm).

If only 1 or 2 diameters are available, the calculation will follow the same rules defined for fibroids volume measured by ultrasound.

Volume of the largest cyst (mL) $= \frac{\pi\times a\times b\times c}{6}$;

a = largest diameter of largest cyst(cm), b = diameter perpendicular to the largest cyst(cm), c = $\frac{a+b}{2}$ (cm)

If only 1 diameter is available, the calculation will follow the same rules defined for fibroids volume measured by ultrasound. If one diameter is recorded as 0 cm, then the value of the diameter will be considered missing and the rule above will apply.

BMD measurements

BMD is assessed using dual-energy X-ray absorptiometry (DEXA). To increase accuracy of the assessment, 2 measurements per location was to be performed at each time point and the mean value will be used for evaluation. If only one scan was performed per location and time point, or one of the two scans is not evaluable, the available valid value will be used as the mean value. Phantom image acquisition is standardized across DEXA imaging facilities according to the instructions specified in an imaging manual. The mean value and percentage change in BMD of each location (the central reading results) from baseline at each time point used for statistical analysis will be calculated by ICON's Medical Imaging group and transferred to Bayer according to the Independent Review Charter.

For BMD baseline, in general, if no Visit 2-repeat scan exists, Visit 2 scan will be used as baseline. Otherwise, Visit 2 -repeat scan will be used as baseline. The BMD baseline visit will be marked in the dataset transferred from ICON's Medical Imaging group.

Repeated measurements at the same visit

If more than one post-randomization measurement is available for a given visit, the first observation will be used in the data summaries if no special reason for the additional observation was provided by the investigator in the CRF, and all observations will be presented in the data listings.

4.6 Blind Review

The results of the final data assessment will be documented in the final list of important deviations, validity findings and assignment to analysis set(s). Any changes to the statistical analysis prompted by the results of the review of study data will be documented in a SAP amendment and, if applicable, in a supplement to this SAP.

5. Analysis Sets

5.1 Assignment of analysis sets

Final decisions regarding the assignment of subjects to analysis sets will be made during the review of study data and documented in the final list of important deviations, validity findings and assignment to analysis set(s) (see Section 4.6).

Full analysis set (FAS): All subjects randomized to vilaprisan treatment groups, excluding randomized subjects who did not start TP 1 due to the study being temporarily paused and all subjects randomized to treatment group B will be included in the FAS. Subjects will be analyzed as randomized.

Safety analysis set (SAF): All subjects randomized to vilaprisan treatment groups who took at least 1 dose of study drug. will be included in the SAF. Subjects will be analyzed as treated.

All subjects randomized to treatment group B will be included in the SAF. Subjects will be analyzed as treated, regardless whether they actually took any medication (i.e., symptomatic, nonhormonal medical treatment).

Modified SAF: will include all subjects in the SAF without any validity findings which are those that may potentially affect the primary safety endpoint. Validity findings leading to exclusion from the modified SAF will be specified in the 'Specification of assessment criteria and identification requirements', which will be finalized before database lock. Subjects will be analyzed as treated.

All safety analyses will be performed on the SAF population. The FAS will be used for the display of all other variables. The analysis for the primary safety endpoint and all BMD related endpoints will be performed on the Modified SAF as well.

Screened subjects who are not included in the FAS will be used solely in disposition tables and will be listed only in Section 16 of the clinical study report, as described elsewhere.

6. Statistical Methodology

6.1 Population characteristics

In general, descriptive statistics will be presented for variables defined in this section. For continuous variables, number of observations, mean, standard deviation, minimum, median, and maximum will be presented. For categorical variables, number and percentage of subjects will be presented. Listings will be provided accordingly.

6.1.1 Subject validity, disposition

Study sample size (all enrolled), subject validity and reasons for excluding subjects from the analyses (all randomized), subject disposition (all enrolled) will be summarized by frequency tables Enrolled means the subject signed the informed consent. The total number and percentage of subjects who prematurely discontinued each epoch with the reasons for discontinuation (all randomized), the number and percentage of subjects who performed adrenal monitoring, skin monitoring, endometrial monitoring and BMD within the safety closeout visit will be provided.

For the overall assessment of disposition, study completion means all phases of the study are completed up to and including the follow-up (FUP) visit.

6.1.2 Demographic, Baseline Characteristics

Demographics and baseline characteristics will be recorded during screening period. As specified in Section 4.5, the measurement closest to the administration of study drug will be used as baseline. All demographic and baseline characteristics will be listed and summarized by treatment and in total according to current existing BHC Global Standard Tables. The descriptive statistics will be presented for the SAF, modified SAF, and FAS.

Demographic and baseline assessments to be summarized will include:

- Age, region/ country and race, ethnicity
- Weight, height, body mass index (BMI)
- Categorized BMI group (<18.5, 18.5-<25, 25-<30, ≥30 kg/m^2)
- Smoking and alcohol consumption
- Baseline BMD of lumbar spine, hip and femoral neck
- Baseline BMD of lumbar spine, hip and femoral neck by scan type (Lunar, Holgic and other)
- Baseline endometrial thickness
- Uterine fibroids visualized at baseline (Yes/No)
- The largest diameter of the largest fibroid, volume of largest fibroid, volume of 3 largest fibroids, location of largest fibroid (measured by ultrasound) at baseline. Baseline uterine volume, endometrial thickness (measured by ultrasound)
- Baseline MBL (by MP) and categorized MBL group (≤80 mL, 80< -150 mL, 150 <- 300 mL, 300 < -500mL and >500 mL)
- Baseline hemoglobin
- Education (years of education, university/colleague degree)

6.1.3 Medical History

Medical history findings (ie, previous diagnoses, diseases or surgeries) will be coded by Medical Dictionary for Regulatory Activities (MedDRA) terms. Number and percentage of subjects with any medical history findings reported will be presented by primary system organ class (SOC) and preferred term (PT) based on the SAF.

All new or worsened findings after signing the informed consent should be documented on the AE eCRF page.

6.1.4 Reproductive and menstrual history

The reproductive history will include information on age at menarche, number of pregnancies, number of births, years since last birth or abortion (at any time and during the last 1 year) and duration of unsuccessful attempts (in months).

For number of pregnancies and number of births, not only the statistics based on continuous variables but also a frequency table based on each specific number recorded in CRF will be presented.

Reproductive history will be summarized for the SAF.

6.1.5 Uterine Fibroids relevant history

The uterine fibroids history includes information on family history, onset of symptoms, diagnosis, and previous medical treatments and procedures, if applicable. Reason for the initial diagnosis, symptoms/finding ever had/currently are collected as well.

Medical treatments relevant for uterine fibroids received ever/ during the last 6 months prior to the study and procedures for uterine fibroids (history, concurrent and planned) will be recorded and presented in this section. Procedures for uterine fibroids (concurrent) will be presented in section of other efficacy variables.

Uterine Fibroids relevant history will be summarized for the SAF.

6.1.6 Heavy menstrual bleeding questions

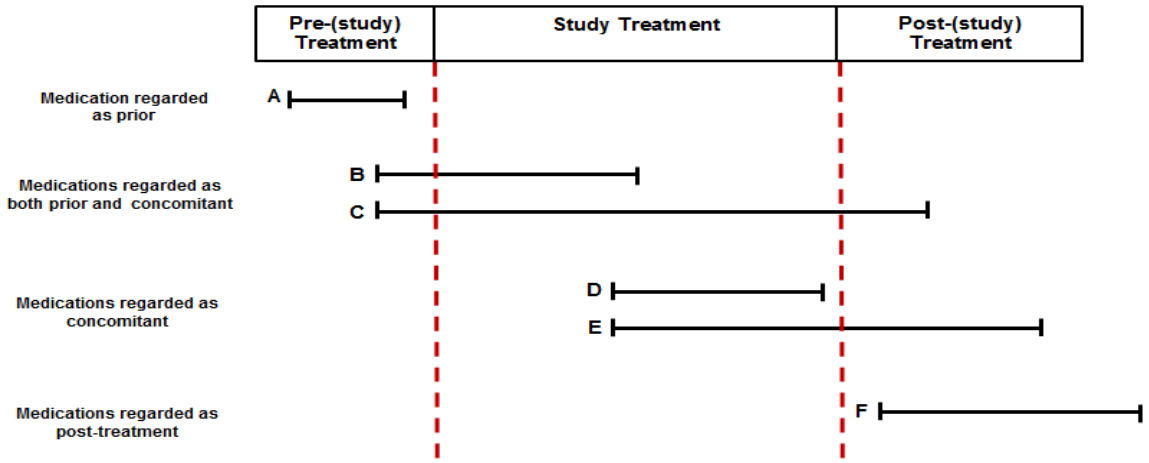
This set of questions has been developed as a tool to identify women with HMB. It could have been used at Visit 1 only and the responses could have been entered directly into the electronic data capturing system (RAVE), which is considered as primary source data. The questionnaire could also have been used as a pre-screening tool. The questions from this questionnaire will be summarized by a frequency table for the SAF.

6.1.7 Prior/Concomitant Medication

For prior/concomitant/post-treatment medications, the following definitions in accordance with the Global Standards Catalogue V4.0 will be used in the analysis (see Figure 6-1):

- Prior medication: Medication taken before start of the study drug intake (regardless of when it ended).
- Concomitant medication: Medication taken during Treatment phase, i.e. between first and last study drug intake (regardless of when it ended). In case it's not clear if a medication is prior, concomitant or post-treatment (for example, due to partially missing dates), medication will be classified as "concomitant".- Post-treatment medication: Medication taken after the Treatment phase, i.e. after last study drug intake.

Figure 6-1: Categories of medication (example)



Categories are prior medication (A, B, C), concomitant medication (B, C, D, E) and post-treatment medication (F). Source: Global Standards Catalogue V4.0

Medication, recorded as prior, concomitant or post-treatment medication in the eCRF, will be coded according to the World Health Organization Drug Dictionary WHODRUG GLOBAL B3 (initially September 1, 2017, within Bayer referred to as '2017SEP'), to the respective Drug Codes with their corresponding Anatomical Therapeutic Chemical (ATC) classification.

The number of subjects taking prior or concomitant medication will be analyzed using frequency tables based on classified data. Analysis of prior and concomitant medication will be done on the SAF.

Nonhormonal medical treatment for symptomatic uterine fibroids at/after the date of randomization for treatment group B will be recorded and coded to ATC codes according to the WHO Drug Dictionary. Subjects in treatment group B who took at least one nonhormonal symptomatic medical treatment will be summarized for prior/concomitant/post-treatment medications separately.

Nonhormonal symptomatic medical treatment：

- Analgesia producing opioids
- Benzodiazepines
- Nonsteroidal anti-inflammatory drugs (NSAIDs)
- Barbiturates and derivatives
- Benzodiazepine derivatives
- Adjunct antidepressant therapy
- Herbal antidepressants
- Monoamine oxidase A inhibitors
- Monoamine oxidase inhibitors, non-selective
- Non-selective monoamine reuptake inhibitors
- Other antidepressants
- Selective serotonin reuptake inhibitors
- Antifibrinolytics
- Blood coagulation factors

- Other haemostatics
- GnRH agonists
- GnRH antagonists
- Adjuvant pain medications
- Analgesic drugs
- Pituitary and hypothalamic hormones and analogues
- Sex hormones and modulators of the genital system
- Anti-inflammatory drugs
- Acetylsalicylic acid and its salts

6.1.8 Exposure and Compliance to Study Treatment

Dose exposure and the compliance for treatment groups A1, A2 and A3 will be analyzed for SAF, modified SAF and FAS. The treatment duration, extent of exposure and compliance will be summarized by treatment group.

6.1.8.1 Treatment duration

Treatment duration is calculated from eCRF page 'Study Drug Exposure.' Missing dates are not replaced, i.e. subjects with missing date of first or last tablet of TP will result in missing treatment duration of respective TP.

Treatment duration per TP (TPx, x=1,…, n) excluding breaks is defined as the number of days from the day of first drug administration until the day of last drug administration during the respective TPs, i.e. Date of last tablet of TPx – Date of first tablet of TPx +1. For treatment group A2, treatment duration excluding breaks by TP will be calculated based on a 168-day period. Overall treatment duration excluding breaks is the sum of available treatment durations of all TPs.

Overall treatment duration including breaks is defined as reference end date – reference start date + 1 without missing data replacement.

Treatment duration for subjects in treatment group B is defined as reference end date – reference start date + 1 without missing data replacement.

6.1.8.2 Extent of exposure

Extent of exposure of each TP is defined as the number of tablets taken during the respective TPs, using information from the Study Drug Accountability eCRF page and eDiary-entry respectively.

Missing number of tablets taken based on the eCRF page will not be replaced, i.e. subjects with any missing number of tablets taken for a pack dispensed based on the eCRF page will result in missing extent of exposure of respective TP. For treatment group A2, extent of exposure by TP will be calculated based on 168-day period. Overall extent of exposure based on the eCRF is the sum of available extent of exposure of all TPs.

Missing number of tablets taken based on eDiary-entry will be replaced with 0. Overall extent of exposure base on eDiary is the sum of the extent of exposure of all treatment periods.

6.1.8.3 Treatment compliance

Treatment compliance of study drug is calculated on the basis of both eCRF (i.e. data collected from drug accountability and study drug exposure pages) and eDiary-entry respectively.

Treatment compliance based on the eCRF by TP is defined as
'$\left(\frac{\text{Extent of exposure in TPx (eCRF)}}{\text{Treatment duration in TPx excluding break(eCRF)}}\right) \times 100\%$', where x=1,…, n.

Overall treatment compliance based on eCRF is calculated as
'$\left(\frac{\text{Overall extent of exposure(eCRF)}}{\text{Overall treatment duration excluding breaks(eCRF)}}\right) \times 100\%$'.

Treatment compliance based on eDiary-entry by TP is defined as
'$\left(\frac{\text{Extent of exposure in TPx (eDiary)}}{\text{Treatment duration in TPx excluding break(eCRF)}}\right) \times 100\%$,' where x=1,…, n.

Overall treatment compliance based on eDiary is calculated as
'$\left(\frac{\text{Overall extent of exposure(eDiary)}}{\text{Overall treatment duration excluding breaks(eCRF)}}\right) \times 100\%$'.

For treatment group A2, treatment compliance by TP will be calculated based on a 168-day period.

Data on duration of treatment, compliance, and used tablets will be presented by listings.

Extent of exposure and treatment compliance for treatment group B will not be analyzed.

6.2 Efficacy

6.2.1 Primary efficacy variable

Not applicable.

6.2.2 Secondary efficacy variable

The analyses for all the secondary efficacy variables will be conducted in the FAS population.

The secondary efficacy variable is number of bleeding days, as defined in Section 4.5.

Number of bleeding days will be shown by treatment groups for the treatment phase only plus 1 or 2 bleeding periods following the end of the last TP depending on treatment regimen.

6.2.3 Other efficacy variables

The analyses for all other efficacy variables will be conducted in the FAS population.

6.2.3.1 Volume of menstrual blood loss related variables

- Volume of menstrual blood loss per 28 days (assessed by MP)
- Volume of menstrual blood loss per bleeding episode (assessed by MP) for first, second, and third bleeding episode after the end of treatment
- Volume of MBL per bleeding episode within the treatment break. All days of the bleeding episodes will be used even if some days of the episode lie outside the treatment break. The first, second and third bleeding episode will be considered. This endpoint is only applicable for VPR groups.
- Amenorrhea (yes/no), defined as menstrual blood loss <2 mL during last 28 days of treatment, based on the MP

- Absence of bleeding (spotting allowed) during the last 28 days of the treatment; based on the UF-DBD

Analysis for volume of menstrual blood loss per 28 days will be done by treatment group and Time interval D (according to Figure 4-1). Amenorrhea and absence of bleeding will be analyzed by treatment group and TP. For treatment group B, volume of menstrual blood loss per bleeding episode (assessed by MP) for first, second, and third bleeding episode after the end of treatment will only be done after the reference end date. Volume of MBL per bleeding episode within the treatment break for treatment group B will not be analyzed.

6.2.3.2 Time to event related variables

- Time to onset of amenorrhea for each treatment period
- Time to onset of controlled bleeding for each treatment period
- Time to start of bleeding after last study drug intake (assessed by the UF-DBD).

Time to event related variables and censoring rules are defined in Section 4.5.

Time to onset variables will be analyzed using the Kaplan-Meier estimates (including number of events, number of censored, 25th, median, and 75th percentiles). This analysis will be performed using the calendar time (days) as unit. Kaplan-Meier plots will also be presented. The analysis will done by treatment group and TP (time intervals B according to Figure 2, excluding screening and FUP). For treatment group A2, TP per 168-day defined in the protocol will be used and shown separately.

Time to start of bleeding after last study drug intake for treatment group B will not be analyzed.

6.2.3.3 Ultrasound examination related variables

- Percent change in volume of largest fibroid compared to baseline (baseline = last value obtained before randomization; measured by ultrasound examination)
- Percent change in volume of 3 largest fibroids compared to baseline (measured by ultrasound)
- Percent change in volume of uterus compared to baseline (measured by ultrasound)
- Percentage of subjects with a volume reduction of ≥25% of the 3 largest fibroids (measured by ultrasound)
- Percentage of subjects with a volume reduction of ≥25% of the largest fibroids (measured by ultrasound)
- Percentage of subjects with a reduction of ≥25% of uterine volume (measured by ultrasound)

Analysis will be done by treatment group and visit/time interval B (according to Figure 4-1).

If multiple volume assessments by ultrasound are available for a time period, the smallest volume will be used for the baseline time period, and the largest volume will be used for any other time periods of Intervals B.

6.2.3.4 Percentage of subjects undergoing surgical treatment

The number and percentage of subjects undergoing surgical treatment will be evaluated in a frequency table by treatment group and by TP and FUP (time interval B). In contrast to other efficacy variables, any fibroid surgeries performed during FUP of subjects who did not start the planned last TP, will be evaluated for FUP rather than for break period.

6.2.3.5 Patient/Clinician-reported outcomes

6.2.3.5.1 Change in UF-DSD individual items compared to baseline

The UF-DSD will be completed daily using the eDiary. The UF-DSD includes 2 items to assess swelling and bloating symptoms using a 5-graded Likert-type severity rating ('no symptom' to 'very severe symptom'), and 2 items using a 0 to 10 numerical rating scale to assess pain at its worst in the abdominal/pelvic and lower back areas with 0 indicating "no pain" and 10 "pain as bad as you can imagine."

Bulk symptoms and pain represent potential domains. Additionally, intake of pain medication is investigated using a 4-point verbal rating scale (no; yes, over the counter (non-prescription) pain medication; yes, prescription pain medication; or yes, both over the counter and prescription pain medication).

Following variables for each domain will be summarized by treatment group and time interval D. For the vilaprisan treatment groups, this will be presented by TP, break, and FUP within treatment group. For the treatment group B, this will be presented by TP and FUP. Line plots with standard deviation (SD) will be generated.

Bulk symptoms (items 1 and 2)

- Mean of subjects' relative frequencies of the scores during baseline (last 28 days before first dose of study drug) and each 28-day period after start of treatment, as well as change from baseline
- Mean of subjects' worst 7 scores during baseline (last 28 days before first dose of study drug) and each 28-day period after start of treatment, as well as change from baseline

Pain (items 3 and 4)

- Mean numerical rating score and absolute change from baseline (last 28 days before first dose of study drug) to each 28-day period after start of treatment
- Mean numerical rating score of the 7 days with worst pain and absolute change from baseline (last 28 days before first dose of study drug) to each 28-day period after start of treatment

For medication intake (item 5), means of proportion of days with "Yes, over the counter (non-prescription) pain medication", "Yes, prescription pain medication" or "Yes, both (over the counter and prescription pain medication)" during baseline (last 28 days before first dose of study drug) and each 28-day period after start of treatment will be summarized, as well as change from baseline.

6.2.3.5.2 Change in Uterine Fibroid Symptom and Quality of Life questionnaire (UFS-QoL) scores compared to baseline

The UFS-QoL consists of 37 items including an 8-item symptom severity scale and 29 health-related quality of life questions, which comprise 6 subscales: concern, activities, energy/mood, control, self-consciousness, and sexual function. Details regarding the calculation of scales are provided in Appendix 1. For each subscale (including symptom severity scale, Revised Activities and Revised HRQoL total score), raw scores and transformed scores will be summarized by treatment group and time interval D, as well as change from baseline. The wort measurement within the time interval will be used for calculations. Largest value for Symptom Severity scale and lowest value for HRQoL scale.Line plots with SD will be generated.

6.2.3.5.3 CGI-I

Clinical Global Impression - Investigator (CGI-I) includes a single item that asks the investigator to describe the subject's overall severity of uterine fibroids symptoms with the response options "none", "very mild", "mild", "moderate", "severe" and "very severe".

CGII will be summarized by a frequency table with the number of observations and percentage by treatment group and visit. Stacked column bar charts will be created by visit.

6.2.3.6 Variables related to laboratory parameters

6.2.3.6.1 Change from baseline in hemoglobin, hematocrit, and ferritin

Hemoglobin, hematocrit, and ferritin will be analyzed in the same way as laboratory measurements described in section other safety variables (see section 6.5.3.6)

6.2.3.6.2 Percentage of subjects with normal hemoglobin >12 g/dL and normal hematocrit >36%

Frequency tables with number of observations and percentage will be displayed by treatment period within treatment group and by Intervals C. A listing will be provided.

6.2.3.6.3 Percentage of subjects with hemoglobin ≤ 10.9 g/dL)

Frequency tables with number of observations and percentage will be displayed by treatment period within treatment group and by Intervals C. A listing will be provided.

6.3 Pharmacokinetics/pharmacodynamics

A separate Pharmacokinetic (PK)/Pharmacodynamic (PD) Evaluation Plan, providing details of the model development and evaluation will be provided before the beginning of the population PK analysis. Evaluation of the data will be presented in a separate PK/PD Evaluation Report.

6.4 Exploratory biomarker analysis

Evaluation of biomarker related analysis will be performed and detailed in a separate document.

6.5 Safety

The safety analysis will be performed on the observed safety data for the safety analysis population.

6.5.1 Primary safety variable

The primary safety variable is the percentage change in BMD of lumbar spine (using the central reading results) from baseline to about one year after SoT. One year is defined as 'Month 12 on treatment' in Table 6-1, i.e. Day 300 after reference start date to Day 420 after reference start date, but no later than the reference end date + 60 days.

The primary safety variable will be analyzed using descriptive statistical methods and two-sided 95% confidence interval (CI) for the mean difference for the percentage change in BMD of lumbar spine from baseline to about one year after SoT will be calculated for each vilaprisan treatment group as compared to treatment group B. Additionally, two-sided 95% CI for each treatment group will be calculated. The primary variable is assumed to be at least approximately normally distributed.

For the percentage change in BMD of lumbar spine from baseline to other time points defined in Table 6-1, the descriptive statistics will also be provided by treatment group.

The percentage change in BMD of lumbar spine from EoT BMD will also be evaluated using summary statistics by treatment group and time interval or month for post treatment period defined in Table 6-1.

The time intervals and month named in Table 6-1 will be presented in tables and figures unless otherwise specified. Month equals 30 days.

Table 6-1: Definition of month and time intervals for BMD summary

	Month	Time interval
Baseline	Baseline	**Baseline** if no Visit 2-repeat scan exists, Visit 2 scan will be used as baseline. Otherwise, Visit 2 -repeat scan will be used. The BMD baseline visit will be marked in the dataset transferred from ICON's Medical Imaging group. (refer to section 4.5)
On treatment period： From reference start date to reference end date + 60 days	Month 1	N/A
	Month 2	If a BMD value performed after the reference start date has been marked as baseline value, it will be analyzed as baseline value and will be excluded from the analysis for on treatment period.
	Month 3	
	Month 4	**Month 6 on treatment** Day 120 after reference start date to Day 240 after the reference start date, but no later than the reference end date + 60 days
	Month 5	
	Month 6	
	Month 7	
	Month 8	
	Month 9	N/A
	Month 10	**Month 12 on treatment** Day 300 after reference start date to Day 420 after reference start date, but no later than the reference end date + 60 days
	Month 11	
	Month 12	
	Month 13	
	Month 14	
	>Month 14	N/A In case of more than one measurement > Month 14, the worst measurement within the time interval will be used for calculations.

	N/A	**EoT BMD** Reference end date ± 60 days If two measurements are available during the EoT BMD time interval, the measurement closest to the reference end date will be use as EoT BMD measurement.
Post treatment period: After reference end date + 60 days	Month 3	N/A
	Month 4	**Month 6 post treatment** Day 120 after the reference end date to Day 240 after the reference end date
	Month 5	
	Month 6	
	Month 7	
	Month 8	
	Month 9	N/A
	Month 10	**Month 12 post treatment** Day 300 after the reference end date to Day 420 after the reference end date
	Month 11	
	Month 12	
	Month 13	
	Month 14	
	>Month 14	N/A In case of more than one measurement > Month 14, the worst measurement within the time interval will be used for calculations.
	N/A	**EoFUP BMD** The last measurement, but no earlier than the reference end date + 60 days

BMD = bone mineral density; EoT=end of treatment; EoFUP= end of follow-up; N/A=not applicable

Percentage change in BMD of lumbar spine from baseline with frequencies and percentages for the following categories: <-6%, ≥-6% - < -3% , ≥-3% - < -1.5% , ≥-1.5% - ≤0%, >0% - ≤1.5%, >1.5% - ≤3% , >3% - ≤6% and >6% will also be summarized by treatment group. The worst outcome per subject during treatment period will be used in the analyses.

Frequency table for the status of BMD of lumbar spine at EoFUP compared with EoT will be provided by treatment group. The categories for the status of BMD of lumbar spine at EoFUP is defined in Table 6-2. Percentage of recovery is defined as change in BMD of lumbar spine from EoT BMD to EoFUP BMD divided by change in BMD of lumbar spine from Baseline to EoT BMD. It will be evaluated in subjects with partially recovered. Number and percentage of subjects for percentage of recovery by categories of ≥ 75 - < 100%, ≥ 50 - <75%, ≥ 25 - <50% and ≥ 0 - <25% will be provided for subjects with partially recovered.

Table 6-2: Status of BMD of lumbar spine at EoFUP

% change in BMD of lumbar spine from baseline to EoT BMD (x)	% change in BMD of lumbar spine from baseline to EoFUP (y)	Status at EoFUP	Percentage of recovery
< 0	≥ 0	Recovered, completely	≥ 100%

< 0	x ≤ y < 0	Recovered, partially	≥ 0% - <100%
< 0	y < x	Not recovered or further decrease	Not applicable
≥ 0	≥ 0	No decrease during both treatment and FUP period	Not applicable
≥ 0	< 0	No decrease during treatment period but worsening during FUP period	Not applicable

Mean-deviation plot for the percentage change in BMD of lumbar spine from baseline will be presented by treatment group for the entire study period (treatment period + post treatment period).

Following listings will be provided separately:

1) subjects that had at least one DEXA scan during the treatment period with either BMD loss ≥6% from baseline to EoT for any location or Z-score ≤ -2 SD for any location.
2) subjects with BMD loss ≥ 3% from baseline for any location or Z-score ≤ -2 SD for any location at EoFUP DEXA scan.

During the study, if the baseline scan at Visit 2 was missing or not evaluable based on the evaluation from the Medical Imaging group of ICON, the subject was requested to have a repeat scan. If it's not possible to have this repeat scan performed before the reference start date, this repeat scan was allowed to be performed shortly after the start of the treatment. Considering the BMD will not change quickly, this can be accepted from a medical point of view. If a BMD value performed after the reference start date has been marked as baseline value, it will be analyzed as baseline value and will be excluded from the analysis for on treatment period defined in Table 6-1.

To make a further investigation, an additional analysis for the primary safety variable with subjects excluded who had their baseline scan later than 1 month after start of treatment will be provided.

6.5.2 Secondary safety variables

6.5.2.1 Endometrial histology (eg, benign endometrium, presence or absence of hyperplasia or malignancy)

For endometrial histology evaluations the following time intervals will be considered:

- Any time point
- Intervals A (according to Figure 4-1)
- Intervals B (according to Figure 4-1)
- Exit examination (according to Figure 4-1)
- Bleeding episode after last TP time windows in subjects with at least 56 days of treatment within the TP (by TP and overall)
    - From start of the subjects´ last treatment period until day 55 of the same TP

- From day 56 of the same treatment period until last day of the1st bleeding episode after TP
- After 1st bleeding episode of the same treatment period, i.e. from 1st day after end of the first bleeding episode to last day of 2nd bleeding episode
- After 2nd bleeding episode of the same treatment period, i.e. from first day after end of 2nd bleeding episode to last day of 3rd bleeding episode
- After 3rd bleeding episode of the same treatment period, i.e. from first day after end of 3rd bleeding episode to last day of 4th bleeding episode

Presentations will be done for

- the safety read (Reader #1 or #5)
- the majority read (Readers #2-4)
- all reads (Readers #1-5)

For the safety read only one result will be available per biopsy (based on either Reader #1 or #5), whereas for the "majority read" and "all reads" more than one reader result needs to be considered.

Analyses will present the number and percentage of either subjects or biopsies. Analyses based on number of subjects will use a worst-case approach within each of the above-mentioned time windows, except "Exit biopsies" where the last available measurement will be used.

Majority read definitions:

Majority read results will only be provided in case biopsy data from Readers #2, #3 and #4 are available.

Majority read will be determined for main results and subcategories (see Table 6-3Table 6-3). First, adequacy (for part II and III) or sufficiency (for part IV) of tissue will be investigated. If at least 2 of the 3 readers consider it adequate/sufficient, the majority for the main results will be determined. For the majority main result, the majority of the respective subcategories will be determined. If no majority result is available (3 different results in 3 readers, 2 different results in 2 readers), either "no consensus" or the worst case will be presented. Table 6-3 presents an overview of all biopsy endpoints including the approach which is used in case no majority is available.

For 'tick all that apply'-subcategories, individual features will be considered the majority in case at least 2 readers have ticked the feature, i.e. for one biopsy more than one subcategory can be the majority.

All read definition:

All read results will be based on all biopsies with results from at least one reader.

A biopsy is considered non-benign in case "Benign endometrium" = "No" or "Endometrial Hyperplasia (2014)" or "Endometrial Hyperplasia (WHO 1994)" or "Malignant Neoplasm" = "Yes".

Worst case across time interval:

Worst case of biopsies or majority results across time windows is determined for each endpoint in Table 6-3Table 6-3.

Table 6-3: Overview of biopsy endpoints including majority result handling and worst case

Part		Endpoint	No majority available	Worst case across time intervals for subject based analyses
Main results				
I	Main diagnosis	Adequate endometrial tissue	- *(not possible)*	"Yes"
II		Benign endometrium Endometrial Hyperplasia (WHO 2014) Malignant Neoplasm	Worst case: List is ordered by severity, from low to high	List is ordered by severity, from low to high
III		Endometrial Polyp	Worst case: yes	"Yes"
IV	PAEC	Tissue sufficient for PAEC read	- *(not possible)*	"Yes"
		PAEC present	Worst case: yes	"Yes"
Subcategories				
II	Main diagnosis	Benign endometrium (select one) Atrophic Inactive Proliferative Disordered Proliferative Secretory including progestin and OCP effect Menstrual Endometritis Other,specify	"no consensus"	- *(will not be presented)*
		Endometrial Hyperplasia (WHO 2014) (select one) Hyperplasia without atypia Atypical hyperplasia / Endometrioid Intraepithelial Neoplasia (EIN)	Worst case: Atypical hyperplasia / Endometrioid Intraepithelial Neoplasia (EIN)	Atypical hyperplasia / Endometrioid Intraepithelial Neoplasia (EIN)
III		Endometrial Polyp (select one) Atrophic Functional Hyperplastic	"no consensus"	- *(will not be presented)*
IV	PAEC	PAEC (tick all that apply) Pre-Decidua Absent Extensive cysts Present Secretory Changes, Extensive Mitoses (required, may be rare) Apoptosis Abnormal Vessels	"no consensus"	- *(will not be presented)*

PAEC = progesterone receptor modulator-associated endometrial changes; WHO = World Health Organization

6.5.2.2 Endometrial thickness

For endometrial thickness the following time intervals will be considered:

- Intervals A (according to Figure 4-1)
- Intervals C (according to Figure 4-1)
- Exit examination

Number of subjects with ultrasound performed will be presented.

Summary statistics for change from baseline in endometrial thickness will be provided. In addition, the proportion of subjects with endometrial thickness > 18 mm will be investigated. A subgroup analysis by exposure group (treatment duration <=90 days, 91 to 180 days, etc.) will present the proportion of subjects with endometrial thickness > 18 mm and cyst like structure either ovary >= 3cm.

To investigate the correlation of endometrial thickness and biopsy results, the proportion of biopsies with hyperplasia without atypia (WHO 2014), atypical hyperplasia / EIN (WHO 2014), PAEC

present, PAEC not present, Extensive cysts present, and Extensive cysts not present will be presented for endometrial thickness > 18mm in ultrasound and endometrial thickness <=18 mm in ultrasound.

6.5.2.3 Percentage change from baseline in BMD measured, hip, and femoral neck by DEXA

The percentage change in BMD of hip, and femoral neck will be summarized in the same way as described for the primary safety variable.

6.5.3 Other safety variables

6.5.3.1 Endometrial histology (diagnosis of PAEC, individual features of PAEC)

Diagnosis of PAEC, individual features of PAEC will be analyzed in the same way as Endometrial histology described in section Secondary safety variables (see section 6.5.2.1)

6.5.3.2 Percentage of subjects with BMD decrease >6% and Z-score < -2 measured at the lumbar spine

Frequency table with number and percentage of subjects with BMD decrease >6% and Z-score <-2 measured at the lumbar spine will be displayed by treatment group during treatment period.

6.5.3.3 Change in Z-score over time

The Z-score and the change in Z-score by BMD location will be summarized by treatment group and time intervals using descriptive statistics.

Z-score with frequencies and percentages for the following categories: <-2, ≥-2 - <-1.5 , ≥-1.5 - <-1 , ≥-1 - ≤0, >0 - ≤1, >1 - ≤1.5, >1.5 - ≤ 2 and >2 will also be summarized by treatment group.

6.5.3.4 Ovarian cysts

Ovarian cysts will be examined by ultrasound. All scheduled and unscheduled assessments will be considered for the analysis. Three analyses will be performed for ovarian cysts:

- Number of subjects showing cyst like structures with largest diameter > 3 cm in the ovary
- Analysis of ovarian cyst episodes (see below for definition)
- Analysis of AEs associated with ovarian cysts

The analysis of **Number of subjects with ovarian cyst like structures > 3 cm in largest diameter** in ovary will be shown by treatment group and by time intervals A and B (according to Figure 4-1). One table will be produced for time intervals A and one table for time intervals B.

For each time interval it will be determined whether an individual subject shows cysts like structures with a largest diameter > 3 cm at least once during that time interval.

In addition to the corresponding time intervals, both tables will contain the assessment for ‘Exit examination.’

Frequencies for the type of cyst like structures identified will be included in the table. Type of cyst like structures will be shown as recorded in the eCRF: ‘Follicle-like structure’, ‘Corpus luteum cyst’, ‘Endometrioma’ and ‘Other’.

The overall assessment for both ovaries will be displayed, and information about ovary laterality will not be included in the table.

If cyst like structures > 3 cm are visualized in the ovaries, unscheduled ultrasound examinations should be performed at least every 4 weeks to document the regression or outcome. An ***ovarian cyst episode*** is defined as the time period from occurrence of a cyst like structure > 3 cm until its resolution. Ovarian cyst episodes will only take account of the occurrence of follicle-like structures and corpus luteum cysts. Two different definitions of "resolution" are used in the analysis:

- the subject does not show in the same ovary (left or right) a cyst like structure of the same type
- the subject does not show in the same ovary (left or right) a cyst like structure of the same type with diameter ≤ 3 cm.

For both definitions, the date of the first observation of the cyst like structure is the **start of the episode**. The date when the cyst like structure appears for the first time as resolved (absence of cyst like structure or cyst like structure with diameter ≤ 3 cm) is considered the **end of the episode**. The length of time between start and end of the episode is the **duration of the episode**.

If the cyst resolution does not occur, then the length of time between the start of the episode and the last observation will be considered the cyst episode duration.

If only the start of the episode is recorded but no follow-up has been performed, then the duration of the episode is unknown.

The analysis will include cysts episodes occurring in both ovaries with no distinction of ovary laterality (left or right). Although the information of ovary laterality will not be displayed in the table, it will be critical to monitor the progression of individual cyst like structures.

The time intervals A and B (according to Figure 4-1) will be considered, therefore two sets of tables will be produced. Ovarian cyst episodes will be assigned to the time interval when the start of the episode occurs.

Two analyses will be produced for ovarian cyst episodes:

- Number of subjects with at least one ovarian cyst episode by time intervals.
- Number of ovarian cyst episodes by time intervals including: i) summary statistics of ovarian cyst episode duration, ii) by categorization of the duration by 4-week intervals, and iii) summary statistics of largest diameter of the cyst like structure per episode and per time interval.

Two tables will be produced for these two analyses, one for each definition of cyst episode resolution.

The analysis of **adverse events related to ovarian cysts** by time intervals will include AEs identified by the Bayer MedDRA query (BMQ) for 'Benign ovarian cyst and associated complications' [V21],], which include the preferred terms (PTs): Haemorrhagic ovarian cyst, Ovarian cyst, Ovarian cyst ruptured and Ovarian cyst torsion.

The table will show number of subjects with AEs, non-serious AE and SAEs. For each one of these AE categories the table will include severity, whether the AE led to drug withdrawal and AE duration. In addition, SAEs will include the reason for being classified as serious as recorded in the eCRF.

The analysis will be shown by treatment group and will be done for pre-treatment AEs, TEAEs and post-treatment AE. AEs will be assigned to the time interval when the AE starts. AEs with onset during a treatment period or up to 8 days after last drug intake in any treatment period will be considered 'on-treatment'. AEs with onset after date of last study drug intake in any treatment period + 8 days and before the start of the next treatment period will be considered 'off-treatment'. The summary table for TEAEs will be presented by overall AEs, on-treatment AEs and off-treatment AEs.

6.5.3.5 AEs

All AEs will be coded by MedDRA terms (current version at time point of analysis) and classified into pre-treatment AEs, treatment-emergent (TEAE) and post-treatment AEs.

All AEs will be flagged as treatment-emergent adverse event (TEAE) except all AEs where there is clear evidence that the event starts before date of first study drug exposure (pre-treatment AEs) or after the date of last study drug exposure + 60 days (post-treatment AE). For Treatment group B, the reference start/end dates (as defined in section 4.5 Data Rules) will be used for this derivation, rather than date of first study drug exposure/ date of last study drug exposure. AEs with clear evidence of start before the date of first study drug exposure are flagged as pre-treatment AE unless the AE worsens at or after the date of first study drug exposure and before or at the date of last study drug exposure + 60 days. AEs starting before but worsening at or after the date of first study drug exposure and before or at the date of last study drug exposure + 60 days will be considered as two events, a pre-treatment AE and a TEAE. The pre-treatment AE will stop at the time of worsening and the new TEAE will start at the time of worsening. Same applies in case an AE starts at or after the date of first study drug administration, but before the date of last study drug administration + 60 days and worsens afterwards then e.g. leading to a TEAE and a post-treatment AE.

Reasons for worsening of an AE are listed below.

- AE intensity/grade is worsened (e.g.: moderate to severe)
- AE changed to a serious event
- AE ends with death
- AE is drug-related

TEAEs will further be categorized into on-treatment and off-treatment events.

All TEAEs should be flagged as on-treatment events except all TEAEs where there is clear evidence that the event starts outside all on-treatment phases. Each on-treatment phase starts with the date of first study drug exposure and ends with the date of last study drug exposure during a treatment period plus 8 days. For subjects in the VPR-6/2 treatment group, TEAEs with an onset within 8 days after end of TP1 (or TP3 or any TP with an odd number), but after start of TP2 (TP4 or after any TP with even number) will be reported as on-treatment for TP2 (TP4 and so on).

TEAEs with clear evidence of start outside all on-treatment phases will be flagged as off-treatment events unless the TEAE worsens during an on-treatment phase. TEAEs starting during off/on-treatment phase, but worsening during an on/off-treatment phase will be considered as two events, an off/on-treatment TEAE and an on/off-treatment TEAE. The off/on-treatment TEAE will stop at the time of worsening and the on/off-treatment TEAE will start at the time of worsening.

Imputation

Missing "Causality to Protocol Procedure" will be imputed with "yes" in case "Causality to study drug" is "no", and with "no" in case "Causality to study drug" is "yes".

Presentation

The tabulation will follow in principle the Bayer Global Standard Tables catalogue (currently V 4.0 default), with the necessary modifications like for on- and off-treatment TEAEs.

Adverse events of special interest (AESIs) will be identified in two different ways

1) The investigators will assess all AEs to determine if they are AEs of special interest (AESIs) and document this in the eCRF. The following AEs are included:
    a. HMB
    b. Liver disorders
    c. Endometrial disorders
    d. Skin disorders
    e. Adrenal disorders
    f. Relevant loss of BMD during treatment
2) Adverse events will be categorized into AE groupings of special interest based on MedDRA groupings as defined within expert statement for compound vilaprisan (Selection of coding conditions applicable for AESIs as defined in the study protocols of vilaprisan phase II/III studies). List of AESIs in current expert statement based on MedDRA Version 22.1 is provided in Appendix 3.

AESIs based on investigator assessment and AESIs based on MedDRA groupings will be presented separately. They will be presented by AESI group (if available), MedDRA primary SOC and PT and study drug. Summary tables for AESIs by maximum intensity and by worst outcome will be produced. Furthermore, listings will be provided.

6.5.3.6 Laboratory parameters

For immunology and coagulation parameters Intervals 'BL/Post baseline' will be used, while for all other laboratory measurements the following time intervals will be considered:

- Intervals C (according to Figure 4-1)

The tabulation will follow in principle the Bayer Global Standard Tables catalogue (currently V 4.0 default), with the necessary modifications like for time interval presentation instead of visit presentation.

The maximum or minimum measurement within the time interval will be used for calculations, following a worst-case approach (for details please refer to Appendix 2). Scheduled as well as unscheduled measurements will be considered.

Central as well as local lab measurements are available. Summary statistics will be provided based on central lab. Local lab measurements will be listed.

Values below the lower limit of quantification (LLOQ) will be set to LLOQ/2.

Values above the upper limit of quantification (ULOQ) will be set to ULOQ.

6.5.3.7 Estradiol

Absolute values will be analyzed using descriptive statistics by treatment group using time intervals C and D. The lowest value (i.e. “worst case”) per subject per time interval will be used in the analyses.

The distribution of estradiol values by treatment group using time intervals C and D like above will be graphically depicted with boxplots. The time course will be shown using line plots depicting the mean and 95% confidence interval (two-sided) and using time intervals C and D as described above. Time intervals including data from less than 10 of randomized subjects in the respective group will be excluded from plots.

For estradiol frequency tables with number and percentage of subjects with estradiol (pg/mL) <13, <30 will be displayed by treatment group and time intervals A and B. For each time interval, the subjects with at least one lower value based on each cut-off in any visit during that time interval will be counted.

6.5.3.8 Cervical smear

Cervical smears findings will be summarized by treatment group and visit with a frequency table.

A listing will be provided with cervical smear results for each subject.

6.5.3.9 Vital signs

Vital signs (heart rate, systolic blood pressure and diastolic blood pressure, and weight) will be summarized by treatment group and visit grouped by TP, including change from baseline where appropriate, using descriptive statistics. Boxplot for absolute values at each visit will be presented by treatment group. The time course by visit and by treatment group will be shown using line plots depicting the mean and 95% CI (two-sided).

For safety closeout visit blood pressure will be measured as triplicates, mean value of triple measurements will be reported in tables.

6.5.3.10 Liver monitoring

Hepatic safety will be evaluated based on the following time intervals:

- BL/Post baseline
- Intervals A

The following parameters will be investigated in addition to the standard lab presentations:

- aspartate aminotransferase (U/L) (AST),
- alanine aminotransferase (U/L) (ALT),
- alkaline phosphatase (U/L) (AP),
- bilirubin in serum (mg/dL).

In case of more than one measurement within a time interval, a worst case approach will be used, i.e. the maximum value will be used for calculations.

Frequency tables presenting number and percentage of subjects fulfilling the following criteria will be produced:

- Normal, >1xULN – <3xULN, ≥3xULN – <5xULN, ≥5xULN – < 8xULN; ≥8xULN - <10xULN; ≥10xULN – <20xULN, ≥20xULN of ALT and AST respectively
- <=1xULN; 1xULN - 2xULN; >=2xULN of Bilirubin
- <=1xULN; 1xULN - 1.5xULN, >=1.5xULN of AP
- ≥ 3×, 5×, 8×, 10×, and 20×ULN of ALT
- ≥ 3×, 5×, 8×, 10×, and 20×ULN of AST
- ≥ 3×, 5×, 8×, 10×, and 20×ULN of ALT or AST
- >1xULN, > 2×ULN of bilirubin
- >1.5 of AP
- >3xULN of ALT or AST accompanied by >1.5xULN Bilirubin (any time point)
- >3xULN of ALT or AST accompanied by >2 xULN in Bilirubin (any time point)

Combined parameters based on Hy's Law criteria

Frequency tables will be provided for

- ≥3×ULN of ALT or AST accompanied by ≥2×ULN of bilirubin (within 30 days afterwards). The event will be assigned to the time interval of the ALT/AST elevation.

Time to event analysis

Kaplan Meier analysis will be performed by time intervals as described above. Time to ALT > 3xULN after start date of the respective time interval will be presented. If no such an increase is observed within the time interval, this is considered a censored observation and the end date of the time interval will be used as date. Tables with number of subjects under risk, cumulative number of ALT >3xULN increases, and estimated probability of increase including 95% confidence intervals within the respective time interval will be presented per 28-day intervals. Furthermore, Kaplan Meier plots will be provided.

Figures and individual subject presentations

An

- eDISH plot (maximum of ALT and AST within time interval × maximum of total bilirubin within 30 days thereafter) and a
- Scatter plot (maximum ALT within time interval x maximum AST within time interval)

will be provided by time interval as described above.

Individual subject presentations

For subjects meeting the following criteria, figures on individual time courses will be presented:

- ≥ 3 x ULN of ALT at any time point

The following lab parameters will be presented: ALT, AST, bilirubin, and AP relative to ULN over time within one figure per subject on absolute scale. It will be indicated within the figure on which days study drug was taken (i.e. start and stop of treatment periods). Furthermore, listings will be provided with laboratory information on liver related parameters, i.e. ALT, AST, bilirubin, AP (relative to upper limit of normal).

6.5.3.11 Adrenal monitoring

Subjects were asked to undergo scheduled adrenal monitoring at the safety closeout visit. The adrenal monitoring comprises the following assessments:

- Adrenal gland imaging
- Laboratory testing
- Signs and symptoms suggestive of hypercortisolism, hyperaldosteronism

All tables generated within this adrenal monitoring section will contain the absolute number of subjects and percentage among subjects who performed the safety closeout visit.

An *overview* table will be provided. This table includes the number and percentage of subjects taking part at the specific adrenal monitoring assessments. The following numbers and percentages of subjects will be provided in addition:

- Subjects with abnormal result in at least one of the assessments.
- Subjects with abnormal result for each specific assessment separately.
- Subjects for whom an external expert recommendation was obtained. Among those, also the number and percentage of subjects with any findings suspicious of an adrenal disorder and recommendation for further handling will be provided.
- Subjects who have seen a local specialist (including information on type of specialist).

Details of the *final diagnosis by the expert panel* will be provided in another table. This table contains the number and percentage of subjects assessed by an expert, subjects needed expert panel meeting, subjects where expert panel met, and subjects with at least one diagnosis by the expert panel. Among those, the following numbers and percentages of subjects will be provided:

- Subjects with agreement of the diagnosis among all three experts.
- Subjects with a specific diagnosis (multiple diagnoses are possible).

A listing of subjects with other diagnosis (i.e. diagnosis entered by the expert panel in a text field) will be provided. A listing of subjects with adrenal tumor will presented as well. It includes diagnosis, time from last study drug until date of panel assessment and possible causal relationship to study drug. Cases where the three experts did not agree on a diagnosis, are presented in a separate listing including also the diagnosis of the local specialist.

A third table will summarize the *diagnosis by a local specialist*. This table contains the number and percentage of subjects with at least one diagnosis by the local specialist. Among those, the numbers and percentages of subjects with a specific diagnosis (multiple diagnoses are possible) will be reported.

A listing of subjects with other diagnosis (i.e. diagnosis entered by a local expert in a text field) will be provided as well as a listing of reasons if a local specialist consultation did not take place despite referral.

A listing of local laboratory measurements will be provided (no descriptive analyses will be performed on those).

A table with *results of the adrenal glands imaging* will be provided. This table contains the number and percentage of subjects who did and who did not perform adrenal imaging (and reasons if not performed), and who have at least one evaluable image. Among subjects having performed adrenal glands imaging the type of procedure (e.g. CT scan with contrast) will be reported. The following numbers and percentages of subjects will be derived among subjects with at least one evaluable image:

- Subjects with one or two evaluable lateralities.
- Subjects with adrenal mass (if yes, in one or in both adrenal glands).
- Subjects with specified diagnostic benign imaging features.
- Subjects with specified indeterminate imaging features.

Reader results where an adjudicator was needed due to discrepant diagnoses between the two readers will be provided in a listing.

6.5.3.12 Skin monitoring

Findings resulting from skin monitoring will be analyzed descriptively by treatment group. All subjects who consent for safety follow-up and attend the Safety closeout visit will undergo a thorough skin examination by a dermatology expert. The information will be recorded on the eCRF assigned to the safety closeout visit. The analysis will show the number of subjects who consented to participate in the Safety Follow-up, how many were assessed by a dermatologist or other clinician, which diagnostic tests were used and the results of the assessment as follows: number of

normal and abnormal findings; and following an abnormal finding which type of abnormality as recorded in the eCRF: 'Precancerous skin lesion', 'Cutaneous sarcoma', 'Other malignant skin tumor' and 'Other skin abnormality'.

Relevant cases will be reported as AEs, potentially as an AESI, and a narrative will be written.

6.5.3.13 UF-DBD bleeding pattern per 28 days and 84 days based on UF-DBD

The bleeding data recorded on the UF-DBD will be summarized using descriptive statistics in the safety analysis set by treatment groups using time intervals C and time intervals D. A reference period of 28 days will be used for most analyses, which corresponds with the time interval D. When a reference period of 84 days is used the time intervals D will be modified accordingly.

The bleeding intensity will be summarized by bleeding category using frequency tables by reference period of 28 days and 84 days.

The maximum bleeding intensity will be summarized using frequency tables by 28 days reference period.

The number of bleeding days will be analyzed by 28 days reference period and with the following bleeding categories:

- non-bleeding,
- spotting-only,
- bleeding (category of bleeding intensity ≥ mild)
- bleeding including spotting (category of bleeding intensity ≥ spotting)
- mild
- moderate
- severe
- very severe

For the safety bleeding analysis, and in contrast with the analysis of efficacy variables based on bleeding data, bleeding induced by endometrial biopsies will not be excluded.

An incomplete period which does not comprise 28 (or 84, respectively) days will be analyzed for any of the considered time periods of Intervals C. The analysis of this incomplete period is based on all subjects for whom the length of the considered time period is not a multiple of 28 (or 84, respectively) days. Subject's data from the last $x < 28$ ($x < 84$) days of the considered time period will be included.

7. Subgroup analyses

Subgroups for endpoints as described in section 6.2.2 and 6.2.3.1

1) Subjects who fulfill the inclusion criteria of HMB > 80.00 mL by MP during screening period (depicted by Baseline Volume of menstrual blood loss per 28 days) and at least one fibroid > 3 cm at baseline by Ultrasound vs. subjects who did not fulfill these criteria at baseline.

8. Document history and changes in the planned statistical analysis

SAP version 1.0. dated 31 Jul 2017

9. References

1. Recording and Evaluation of Bleeding Data Version 6.0 (BHC-RD-OI-107)
2. Nikica Živković1, Krešimir Živković2, etal. Measuring the volume of uterine fibroids using 2- and 3- dimensional ultrasound and comparison with histopathology. Acta Clin Croat. 2012; 51:579-589

Appendix 1: Handling of Patient-report outcomes

UFS-QoL scores

To calculate a symptom score for symptom severity, create a summed score from the items listed below and then use the formula below the table to transform the value. This will provide symptom scores where higher score values are indicative of greater symptom severity or bother and lower scores will indicate minimal symptom severity (high scores= bad).

Table 9-1 Formula for calculating symptom score for symptom severity

Scale	Sum Item Values	Lowest and Highest Possible Raw Scores	Possible Raw Score Range
Symptom Severity	Sum 1-8	8, 40	32

Transformation for Symptom Severity raw scores ONLY:

$$\text{Transformed Score} = \frac{(Actual\ raw\ score - lowest\ possible\ raw\ score)}{Possible\ raw\ score\ range} \times 100\%$$

For the HRQoL subscales (concern, activities, energy/mood, control, self-conscious, and sexual function), create summed scores of the items listed below for each individual subscale. To calculate the HRQoL total score, sum the value of each individual subscale (do not sum individual items). Use the formula below the table to transform all values. Higher scores will be indicative of better HRQoL (high = good).

Table 9-2 Formula for calculating HRQoL subscales and total score

Scale	Sum Item Values	Lowest and Highest Possible Raw Scores	Possible Raw Score Range
Concern	9+15+22+28+32	5,25	20
Activities	10+11+13+19+20+27+29	7,35	28

Energy/mood	12+17+23+24+25+31+35	7,35	28
Control	14+16+26+30+34	5,25	20
Self-conscious	18+21+33	3, 15	12
Sexual function	36+37	2, 10	8
HRQoLTOTAL	Sum of 6 Subscale Scores	29, 145	116
Revised activities	11+13+19+20+27	5,25	20
Revised HRQLTOTAL	Sum of 6 Subscale Scores (replacing activities with revised activities)	27,135	108

Formula for transformation of HRQoL raw scores ONLY:

$$\text{Transformed Score} = \frac{(Highest\ possible\ score - Actual\ raw\ score)}{Possible\ raw\ score\ range} \times 100\%$$Missing Items

For the subscale analyses, if< 50% of the scale items are missing, the scale should be retained with the mean scale score of the items present used to impute a score for the missing items. If ≥ 50% of the items are missing, no scale score should be calculated, the subscale score should be considered missing. If a subscale score is missing, the HRQoL total cannot be calculated.

Appendix 2 : The maximum or minimum measurement for laboratory parameters

Table 9-3 The maximum or minimum measurement for laboratory parameters

Laboratory category	Laboratory test	lowest	highest
VITAMINS	25-Hydroxyvitamin D	X	
COAGULATION	Activated Partial Thromboplastin Time		X
HORMONES	Adrenocorticotropic Hormone	X	X
GENERAL CHEMISTRY	Alanine Aminotransferase		X
GENERAL CHEMISTRY	Albumin	X	
GENERAL CHEMISTRY	Alkaline Phosphatase		X
IMMUNOLOGY	Alpha-1 Antitrypsin		X
IMMUNOLOGY	Anti Mitochondiral Antibody		X
IMMUNOLOGY	Antinuclear Antibodies		X
GENERAL CHEMISTRY	Aspartate Aminotransferase		X
HEMATOLOGY	Basophils	X	X
HEMATOLOGY	Basophils/Leukocytes		X
GENERAL CHEMISTRY	Bilirubin		X
URINALYSIS	Bilirubin		X
URINALYSIS	Blood		X
GENERAL CHEMISTRY	Bone Specific Alkaline Phosphatase		X
MICROBIOLOGY	Brucellae		X
GENERAL CHEMISTRY	Calcium	X	X
PROTEINS	Ceruloplasmin		X
GENERAL CHEMISTRY	Chloride		X
GENERAL CHEMISTRY	Cholesterol		X
	Choriogonadotropin Beta		X
HORMONES	Choriogonadotropin Beta		X

URINALYSIS	Choriogonadotropin Beta		X
HORMONES	Cortisol	X	X
GENERAL CHEMISTRY	Creatine Kinase		X
GENERAL CHEMISTRY	Creatinine		X
IMMUNOLOGY	Cytomegalovirus DNA		X
GENERAL CHEMISTRY	Direct Bilirubin		X
HEMATOLOGY	Eosinophils		X
HEMATOLOGY	Eosinophils/Leukocytes		X
IMMUNOLOGY	Epstein-Barr Capsid IgG Antibody		X
IMMUNOLOGY	Epstein-Barr Capsid IgM Antibody		X
IMMUNOLOGY	Epstein-Barr Nuclear Antigen 1 IgG Ab		X
HEMATOLOGY	Ery. Mean Corpuscular Hemoglobin		X
HEMATOLOGY	Ery. Mean Corpuscular Volume	X	
HEMATOLOGY	Erythrocytes	X	
HORMONES	Estradiol	X	
GENERAL CHEMISTRY	Ferritin	X	
HORMONES	Follicle Stimulating Hormone		X
GENERAL CHEMISTRY	Gamma Glutamyl Transferase		X
URINALYSIS	Glucose		X
IMMUNOLOGY	HCV PCR Viral Load		X
GENERAL CHEMISTRY	HDL Cholesterol	X	X
HEMATOLOGY	Hematocrit		X
HEMATOLOGY	Hemoglobin	X	
GENERAL CHEMISTRY	Hemoglobin A1C		X
IMMUNOLOGY	Hepatitis A Virus Antibody		X
IMMUNOLOGY	Hepatitis A Virus Antibody IgM		X
IMMUNOLOGY	Hepatitis B Virus Core Antibody		X
IMMUNOLOGY	Hepatitis B Virus DNA		X
IMMUNOLOGY	Hepatitis B Virus Surface Antibody		X
IMMUNOLOGY	Hepatitis B Virus Surface Antigen		X
IMMUNOLOGY	Hepatitis C Virus Antibody Surface		X
IMMUNOLOGY	Hepatitis D Virus Antibody		X
IMMUNOLOGY	Hepatitis D Virus IgG Antibody		X
IMMUNOLOGY	Hepatitis D Virus IgM Antibody		X
IMMUNOLOGY	Hepatitis D Virus RNA		X
IMMUNOLOGY	Hepatitis E Virus Antibody		X
IMMUNOLOGY	Hepatitis E Virus IgG Antibody		X
IMMUNOLOGY	Hepatitis E Virus IgM Antibody		X
IMMUNOLOGY	Hepatitis E Virus RNA		X
IMMUNOLOGY	Herpes Simplex Virus 1 IgG Antibody		X
IMMUNOLOGY	Herpes Simplex Virus 1 IgM Antibody		X
IMMUNOLOGY	Immunoglobulin A		X
IMMUNOLOGY	Immunoglobulin G		X
IMMUNOLOGY	Immunoglobulin M		X
GENERAL CHEMISTRY	Iron	X	
URINALYSIS	Ketones		X
GENERAL CHEMISTRY	Lactate Dehydrogenase		X
GENERAL CHEMISTRY	LDL Cholesterol		X
MICROBIOLOGY	Leptospirae		X
HEMATOLOGY	Leukocytes	X	
URINALYSIS	Leukocytes		X
HORMONES	Luteinizing Hormone		X
HEMATOLOGY	Lymphocytes	X	
HEMATOLOGY	Lymphocytes/Leukocytes		X
HEMATOLOGY	Monocytes	X	
HEMATOLOGY	Monocytes/Leukocytes		X
IMMUNOLOGY	MPO ANCA		X
HEMATOLOGY	Neutrophils	X	
HEMATOLOGY	Neutrophils/Leukocytes		X
URINALYSIS	Nitrite		X
GENERAL CHEMISTRY	Osteocalcin		X

HEMATOLOGY	Platelets	X	
GENERAL CHEMISTRY	Potassium	X	x
IMMUNOLOGY	PR3 ANCA		X
HORMONES	Progesterone	X	
HORMONES	Prolactin		X
GENERAL CHEMISTRY	Protein		X
URINALYSIS	Protein		X
COAGULATION	Prothrombin Intl. Normalized Ratio	X	
COAGULATION	Prothrombin Time		X
GENERAL CHEMISTRY	Pseudocholinesterase		X
IMMUNOLOGY	Smooth Muscle Antibody		X
GENERAL CHEMISTRY	Sodium	X	
HORMONES	Testosterone		X
HORMONES	Thyrotropin	X	X
GENERAL CHEMISTRY	Time to event		X
GENERAL CHEMISTRY	Total Iron Binding Capacity		X
IMMUNOLOGY	Toxoplasma gondii antibody		X
GENERAL CHEMISTRY	Triglycerides		X
GENERAL CHEMISTRY	Type I Collagen C-Telopeptides		X
URINALYSIS	Urobilinogen		X
	all parameters not otherwise specified		X

Appendix 3 : List of AESIs in current expert statement based on MedDRA Version 22.1 (version 5.0, status 2020-02-19)

Table 9-4 List of AESIs

#	AESI	Grouping
1	HMB (especially after end of treatment): HMB will be documented in detail throughout the study	**MTG:** [BMQ] Increased female genital bleeding[V8]
2	Progesterone receptor modulator-associated endometrial changes (PAEC) (not an AESI in 16953)	**PT selection:** PT: Endometrial thickening PT: Endometrial disorder PT: Endometrial hypertrophy PT: Biopsy endometrium abnormal PT: Biopsy uterus abnormal
3	Ovarian cysts (), not an AESI in study 16953)	**MTG:** [BMQ] Benign ovarian cyst and associated complications[V21]
4	Liver enzymes	**SMQ**: Drug related hepatic disorders - comprehensive search (SMQ)[V29]
5	Endometrial hyperplasia (all subcategories according to WHO 2014 (and WHO 1994) classification	**PT selection:** PT: Endometrial hyperplasia PT: Biopsy endometrium abnormal PT: Endometrial dysplasia PT: Endometrial metaplasia

		PT: Biopsy uterus abnormal
6	Endometrial thickening >18 mm	**PT selection:** PT: Endometrial thickening PT: Endometrial disorder PT: Endometrial hypertrophy
7	Loss of BMD	**SMQ:** Osteoporosis/osteopenia (SMQ) (V21)
8	Adrenal neoplasms, Adrenal diagnoses related to cortical hyperfunction and cortical hypofunction	**PBMQ:** Adrenocortical gland pathology including tumors and related hormonal changes and conditions vilaprisan(narrow scope) PT Addison's disease PT Adrenal adenoma PT Adrenal androgen deficiency PT Adrenal atrophy PT Adrenal calcification PT Adrenal cortex dysplasia PT Adrenal cortex necrosis PT Adrenal cyst PT Adrenal disorder PT Adrenal gland abscess PT Adrenal gland cancer PT Adrenal gland cancer metastatic PT Adrenal gland injury PT Adrenal gland tuberculosis PT Adrenal glomerular zone abnormal PT Adrenal haematoma PT Adrenal haemorrhage PT Adrenal insufficiency PT Adrenal insufficiency neonatal PT Adrenal mass PT Adrenal neoplasm PT Adrenal suppression PT Adrenal thrombosis PT Adrenalitis PT Adrenocortical carcinoma PT Adrenocortical insufficiency acute PT Adrenocortical insufficiency neonatal PT Adrenogenital syndrome PT Adrenoleukodystrophy PT Adrenomegaly PT Benign neoplasm of adrenal gland PT Catecholamine crisis PT Congenital adrenal gland hypoplasia PT Congenital anomaly of adrenal gland PT Cortisol deficiency PT Cushingoid PT Cushing's syndrome PT Familial glucocorticoid deficiency PT Glucocorticoid deficiency PT Haemorrhagic adrenal infarction PT Hyperadrenalism PT Hyperadrenocorticism PT Hyperaldosteronism PT Hypercorticoidism PT Hypoaldosteronism PT Metastases to adrenals PT Mineralocorticoid deficiency

		PT Myelolipoma PT Primary adrenal insufficiency PT Primary hyperaldosteronism PT Pseudohypoaldosteronism PT Secondary adrenocortical insufficiency PT Secondary aldosteronism PT Steroid withdrawal syndrome PT Triple A syndrome
9	Cutaneous sarcomas	**PT selection:** PT: Dermatofibrosarcoma protuberans PT: Atypical fibroxanthoma PT: Leiomyosarcoma PT: Liposarcoma PT: Skin angiosarcoma PT: Kaposi´s sarcoma PT: Malignant fibrous histiocytoma PT: Soft tissue sarcoma

Appendix 4: Mapping of MP icons to MBL volume (mL)

Sanitary product	Menstrual spot shape	Menstrual blood loss (mL)
always Ultra normal	Towel pictogram a	0.79
	Towel pictogram b	1.51
	Towel pictogram c	3.08
	Towel pictogram d	4.64
	Towel pictogram e	6.75
	Towel pictogram f	11.91
always Ultra night	Towel pictogram a	0.74
	Towel pictogram b	2.2
	Towel pictogram c	4.21
	Towel pictogram d	6.84
	Towel pictogram e	9.42
	Towel pictogram f	15.2
always Ultra long	Towel pictogram a	1.36
	Towel pictogram b	1.89
	Towel pictogram c	3.68
	Towel pictogram d	6.12
	Towel pictogram e	8.69
	Towel pictogram f	13.41
Other Towel	Towel pictogram a	2.29
	Towel pictogram b	2.84
	Towel pictogram c	3.29
	Towel pictogram d	7.46
	Towel pictogram e	15.18
	Towel pictogram f	20.12
o.b. Pro Comfort light	Tampon pictogram a	0.16
	Tampon pictogram b	0.55
	Tampon pictogram c	1.19
	Tampon pictogram d	2.98
o.b. Pro Comfort regular	Tampon pictogram a	0.35
	Tampon pictogram b	1.31
	Tampon pictogram c	2.09
	Tampon pictogram d	5.66

Sanitary product	Menstrual spot shape	Menstrual blood loss (mL)
o.b. Pro Comfort super	Tampon pictogram a	1.29
	Tampon pictogram b	1.87
	Tampon pictogram c	3.49
	Tampon pictogram d	9.37
o.b. Pro Comfort super plus	Tampon pictogram a	1.34
	Tampon pictogram b	3.36
	Tampon pictogram c	5.81
	Tampon pictogram d	12.23
Tampax light	Tampon pictogram a	1.24
	Tampon pictogram b	1.52
	Tampon pictogram c	1.73
	Tampon pictogram d	3.1
Tampax regular	Tampon pictogram a	0.63
	Tampon pictogram b	1.58
	Tampon pictogram c	2.2
	Tampon pictogram d	5.62
Tampax super	Tampon pictogram a	1.29
	Tampon pictogram b	2.38
	Tampon pictogram c	4.01
	Tampon pictogram d	7.76
Tampax super plus	Tampon pictogram a	1.2
	Tampon pictogram b	2.56
	Tampon pictogram c	4.98
	Tampon pictogram d	9.49
Other Tampon	Tampon pictogram a	3.55
	Tampon pictogram b	4.33
	Tampon pictogram c	6.25
	Tampon pictogram d	8.3