SAP for EHDEN HMB

Martin Lavallee

2023-07-25

Table of contents

[1. Background 1](#_Toc141179782)

[2. Research Questions 2](#_Toc141179783)

[3. Objectives 2](#_Toc141179784)

[3.1 Primary Objectives 2](#_Toc141179785)

[4. Study Design 2](#_Toc141179786)

[4.1 Study Population 3](#_Toc141179787)

[4.2 Exposure Definition 3](#_Toc141179788)

[4.3 Outcome Definition 3](#_Toc141179789)

[5. Data Sources 3](#_Toc141179790)

[6. Analysis Plan 4](#_Toc141179791)

[6.1 Cohort Diagnostics 4](#_Toc141179792)

[6.2 Baseline Characteristics 4](#_Toc141179793)

[6.3 Treatment Landscape 5](#_Toc141179794)

[6.3.1 Prevalence of Drug Exposures 6](#_Toc141179795)

[6.3.2 Constructing the treatment history 7](#_Toc141179796)

[6.3.3 Treatment Pathways 7](#_Toc141179797)

[6.3.4 Duration of Drug use 7](#_Toc141179798)

[6.4 Procedure use post-index 7](#_Toc141179799)

[6.5 Incidence Analysis 8](#_Toc141179800)

# 1. Background

This study is a continuation of a Bayer heavy-menstrual bleed (HMB) study that described first, the incidence rate of HMB; two, describe clinical characteristics of women diagnosed with HMB; and third describe the treatment pathways in five healthcare databases covering the UK, Germany, France and the US from the time period of 2000 to 2022. The first part of this study was exclusively an internal study within Bayer. The second part of the HMB study has similar objectives to part 1 with the key difference being a focus on European databases (UK CPRD) and public expansion of the analysis to data providers involved in the European databases in the European Health Data and Evidence Network (EHDEN). The purpose of this document is elaborate on the study design constructed by Odysseus Data Services. A separate protocol has been written by Bayer for HMB part 2.

# 2. Research Questions

* What are the characteristics of women diagnosed with HMB in terms of demographics, comorbidities, procedures, and comedication?
* What are the treatment patterns for HMB?
* What is the incidence of HMB across different countries and data sources?

# 3. Objectives

## 3.1 Primary Objectives

The primary objectives of this study are to:

* Describe demographics and baseline clinical characteristics of women of reproductive age diagnosed with HMB between 2000 and 2022 across a network of European healthcare databases.
* Describe treatment utilization and longitudinal treatment pathways of women of reproductive age diagnosed with HMB across different countries and data sources between the years 2000 and 2022 across a network of European healthcare databases.
* Estimate incidence of HMB among women between the ages of 11 to 55 between the years 2000 and 2022 across a network of European healthcare databases.

# 4. Study Design

This large-scale retrospective cohort study will be conducted across a network of European healthcare databases standardized to the OMOP CDM. Databases assessed in this study are either: a) Bayer data assests (CPRD GOLD and AURUM) or b) participating EHDEN data-partners. The final list of EHDEN data partners will be provided in the Data Sources section.

The study period begins on January 1, 1999 (or earliest date of data availability following this date) and end at the latest date of data availability in each database. Persons involved in this study must have a minimum of one-year prior observation in the database, therefor the indexing period begins on January 1, 2000.

The study population for this study are women between the ages of 11 and 55 who have a first-time diagnosis of HMB, given that they have at least 365 days of continuous observation prior to the index date. More details of the study population are found in the study population section.

## 4.1 Study Population

The study population are women between the ages of 11 to 55 who with a diagnosis heavy-menstrual bleeding (index event) between the years 2000 and 2022. Women entering the HMB cohort must have a minimum of 365 days of prior observation. Female patients are excluded from the cohort if they have one of the following criteria:

1. Observation of a hysterectomy or bilateral ovariectomy
2. Observation of a menopause diagnosis
3. Observation of other gynecological bleeding (vaginal bleeding outside the menstrual cycle)

Female patients are followed up until end of continuous observation or observation of one for the follow, whichever occurs first:

1. Death
2. Turning age 55 (indicating transition into natural menopause)
3. Observation of a hysterectomy or bilateral ovariectomy
4. Observation of a menopause diagnosis

## 4.2 Exposure Definition

There is no exposure definition in this study.

## 4.3 Outcome Definition

There is no outcome definition in this study.

# 5. Data Sources

The datasources for this study include: a) Bayer OMOP assets and b) European databases that are part of EHDEN. All databases used in this study have been standardized to the OMOP CDM.

Table 1: Data Sources used in study

| Name | Country | Type | Size | Availability | Description |
| --- | --- | --- | --- | --- | --- |
| CPRD AURM | UK | EHR |  |  | Primary care data from a participating electronic system in England |
| CPRD AURM | UK | EHR |  |  | Primary care data from across the UK |
| Hospital del Mar | Spain | EHR |  |  | Hospital data from Barcelona Spain |

The above table will be expanded upon receiving information of the EHDEN participants.

# 6. Analysis Plan

## 6.1 Cohort Diagnostics

Prior to running any specific analysis, we will evaluate the HMB cohort using the OHDSI R package CohortDiagnostics. This package produces metrics such as cohort counts in the database, incidence rates (by calendar year, age and gender), time distributions, cohort attrition and breakdown of index events. Evaluation of these metrics helps ensure that the clinical cohort is indeed reliable in capturing HMB in the OMOP database.

## 6.2 Baseline Characteristics

We assess baseline characteristics based on an observation window of 365 to 1 days prior to the index date. Categorical covariates are reported using the count and percentage. Continuous covariates are reported using the median, 25th and 75th percentile.

**Demographics**

* Age at HMB diagnosis as 5-year categories
* Age at HMB diagnosis as continuous
* Race (if available)
* Ethnicity (if available)
* Year of HMB Diagnosis (per calendar year)

**Comorbidity Scores**

* BMI (if available) continuous and in following categories
  + Underweight (< 18.5 kg/m^2)
  + Normal Weight (18.5 - 24.9 kg/m^2)
  + Overweight (25.0 - 29.9 kg/m^2)
  + Obese (30.0 - 39.9 kg/m^2)
  + Morbidly obese (> 40.0 kg/m^2)
* Charlston Comorbidity Score (as a continuous score)
* CHADs2Vasc (as a continuous score)

**Concept-based**

* Drug Era individual and rolled up to ATC2 Categories
* Condition Era individual and rolled up to ICD10 Chapters

**Note**: Concept-based covariates are based on prevalence concepts accumulated via FeatureExtraction. .

**Cohort-based**

* Conditions
  + Underlying Causes of HMB (individual)
    - Uterine endometrial polyps
    - Adenomyosis
    - Uterine leiomyoma (fibroids) (subserous or intramural)
    - Uterine malignancy and endometrial hyperplasia
    - Coagulopathy / coagulation disorders
    - Ovulatory dysfunction (including hypothyroidism, polycystic ovary syndrome, adrenal disorders, hyperprolactinemia, hypothalamic disorders)
    - Endometrial dysfunction
    - Iatrogenic HMB
    - Endometriosis
    - Idiopathic menorrhagia
  + Other Conditions
    - Diabetes Mellitus
    - Polycystic Ovary Syndrome (PCOS)
    - Dysmenorrhea
    - Pain
    - Anemia
    - Iron deficiency anemia
* Drugs
  + Antithrombotic Agents
  + Antidepressants
  + Tamoxifen
  + Antipsychotics
  + Gonadal Steroids
* Procedures
  + Endometrial ablation
  + Uterine artery embolization (UAE)
  + Myomectomy
  + Blood transfusion

## 6.3 Treatment Landscape

In this study we assess the landscape of treatment for female patients with HMB. The term treatment landscape encapsulates three metrics of interest regarding drug utilization post-index: first, the prevalence of drug use at certain time windows post-index; second, the patterns of treatments (called treatment pathways) including single drugs and combinations of treatment; and third, the length of time spent taking the drug exposures of interest, both as single exposures or in combination. The drug exposures of interest are:

* Tranexamic acid (B02AA02)
* Progestin only regimens:
  + Medroxyprogesterone acetate (MPA) (G03AC06)
  + Oral norethindrone acetate (NETA) (G03DC02)
  + Desogestrel (G03AC09)
  + Etonogestrel implant (G03AC08)
* Non-steroidal anti-inflammatory drugs (NSAIDs) (M01A)
* Combined oral hormonal contraceptives
  + Dienogest and estradiol (G03AB08; sequential combinations)
  + Nomegestrol and estradiol (G03AA14; fixed combinations)
* Selective progesteron receptor modulators (G03AD02, G03XB02)
* Danazol (G03XA01)
* Gonadotropin releasing hormone analogues (L02AE)
* Intrauterine devices (G02BA)
* Iron preparations (B03A)

For the treatment landscape analysis, we must define a set of drug exposure cohorts where the index event of each cohort is prescription to one of the list drugs of interest between the years of 2000 and 2022. People remain in the cohort if they are females between the ages of 11 to 55 and have at least 365 days of prior observation. A persons time in the cohort prior to exit is determined by a drug era. The drug era is the span of time a person is considered continuously exposed to the same treatment. Eras are built by binding successive drug exposure events into a single duration where the person is inferred to be continuing use of that same treatment. We allow for a maximum of 30 days between drug exposure records to consider them part of the same era. Thus if a person has two of the same drug exposures within 30 days of each other, they are bundled into one era where the time elapsed is marked by the date of the successive exposure. This also means that if there is a gap greater than 30 days between the same treatment, this would indicate two separate eras in the treatment history for the individual person. A person is observed in the drug exposure cohort until either the end of a continuous era or one of the following censoring events, whichever occurs first:

1. Death
2. Turning age 55 (indicating transition into natural menopause)
3. Observation of a hysterectomy or bilateral ovariectomy
4. Observation of a menopause diagnosis

### 6.3.1 Prevalence of Drug Exposures

The first metric in our treatment landscape analysis is the prevalence of drug exposure over a set of defined time intervals. Prevalence is defined as the proportion of the at-risk population who used the medications of interest during the interval. The time windows for prevalence of drug exposures for women in the HMB cohort are: 0 days, 1 to 183 days, 184 to 365 days, 366 to 730 days, and 731 to 1825 days post index. We report the count and percentage for each drug exposure of interest (from the listed above) at each time window.

### 6.3.2 Constructing the treatment history

Before continuing our treatment landscape analysis, we must align the drug exposure cohorts into a treatment history where we define when there is a discontinuation of a treatment, combination or a switch in treatments. A treatment discontinuation is defined when there is a gap of at least 30 days between successive exposures of the same treatment. A treatment combination is defined as an overlap in two different exposures of at least 30 days. This means if two drugs are being taken by the same patient for at least 30 days it is considered a treatment combination. If the overlap of two drugs is less than 30 days it is considered a switch to the drug whose era persists. For example if drug A starts first and then overlaps with drug B for 15 days, with drug B continuing onward, the treatment history would consider this a switch from drug A to drug B and no combination.

Following the alignment of the treatment history for each patient, we can enumerate each unique treatment pathways identified post index and determine the duration spent on each treatment. We treatment the duration as separate for a combination of treatments to that of an individual drug since we consider the start of a combination to be a switch from a single drug to a combination.

### 6.3.3 Treatment Pathways

For each female with HMB, we determine the sequence of treatments taken over the patient history. Next we enumerate each unique treatment sequence identified in the population. This unique sequence is considered a pathway. Treatment pathways with less than 20 patients identified are dropped from the analysis. The treatment sequence results will be displayed using a sankey diagram, complemented by a summary table that shows the enumeration of each specific sequence.

### 6.3.4 Duration of Drug use

Using the treatment history, we calculate the median (and 95% confidence interval) time spent on a drug exposure using the Kaplan-Meier (KM) methodology. We consider a combination of exposure to be a separate calculation to individual exposures in this analysis.

## 6.4 Procedure use post-index

In this analysis we are also interested in procedures used after the HMB diagnosis between the years of 2000 and 2022. The procedures of interest are:

* Hysterectomy
* Endometrial ablation
* Uterine artery embolization (UAE)
* Myomectomy

For the procedure analysis we define cohorts where the index event for the procedure cohort is a first time occurrence of one of the procedures of interest in the patient history. Women in the procedure cohort must have at least 365 days of prior observation. Women exit the cohort on the day of the procedure occurrence.

The first metric in our analysis of procedures is prevalence of over a set of defined time intervals. The time windows for prevalence of procedures for women in the HMB cohort are: 0 days, 1 to 183 days, 184 to 365 days, 366 to 730 days, and 731 to 1825 days post index. We report the count and percentage for each procedure of interest (from the listed above) at each time window.

The second metric in our analysis of procedures is duration of time until the procedure intervention. We define the time to intervention as the duration between the index date of the HMB cohort to the index date of the procedure exposure cohort. We calculate the median (and 95% confidence interval) time spent until procedure using the Kaplan-Meier (KM) methodology.

## 6.5 Incidence Analysis

It is of interest to ascertain the overall and yearly incidence of HMB in each participating EU databases. We define the outcome for this incidence analysis to the target population defined in [Section 4.1](#sec-target), women between the ages of 11 to 55 with first time diagnosis of HMB. The denominator for this incidence calculation will be women age 11 to 55 in the database who exit the cohort based on whether they have experienced one of the following events: turn age 55, observation of a menopause diagnosis, observation of a hysterectomy or observation of a bilateral ovariectomy. Incidence is defined based on the following formula:

The incidence of HMB will be reported overall and per calendar year to show any variation over time. The incidence will be calculated using the CohortIncidence package from HADES.