SAP for EHDEN HMB

Martin Lavallee

2023-06-27

Table of contents

# 1. Background

This study is a continuation of a Bayer Heavy-menstrual bleed (HMB) study, which had been conducted internally. The second part of the HMB study sees it expand to the European Health Data and Evidence Network (EHDEN) aiming to characterize HMB across a network of European database who have had their data standardized to the OMOP CDM. Part 2 of the HMB study has a dedicated protocol, the purpose of this document is provide further details of part 2 of the analysis as supported by Odysseus Data Services.

# 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.
* Describe treatment utilization and longitudinal treatment pathways of women of reproductive age diagnosed with HMB across different countries and data sources.
* Estimate incidence of HMB among women between the ages of 11 to 55.

# 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 will begin 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 will begin 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 populaiton section.

## 4.1 Study Population (Target Cohort)

The target cohort for this study are women who have a first-time diagnosis of HMB, given that they have at least 365 days of continuous observation prior to the index date. We will refer to this as the HMB cohort. This target cohort serves as a basis for estimating the number of women with incident heavy-menstrual bleed in the participating databases. Women are included in this cohort if they satisfy the following criteria:

1. Women is between the ages of 11 and 55 years old
2. No prior observation of a hysterectomy or bilateral ovariectomy
3. No prior observation of menopause based on a diagnosis code
4. No prior observation of other gynecological bleeding

Women exit from the HMB cohort at the end of continuous observation in the database or if one of the following censoring criteria occur:

* Death
* at least one visit occurrence over the age of 55 (indicating they are 55 years-old)
* at least one observation of a menopause diagnosis
* at least one observation of a hysterectomy procedure
* at least one observation of a bilateral ovariectomy procedure

## 4.2 Exposure Definition

There is no exposure definition in this study.

## 4.3 Outcome Definition

### 4.3.1 Heavy-mentrual Bleed

It is of interest to ascertain the incidence of HMB in the participating databases of this study. We define the outcome for this incidence analysis to be the same as that defined in the target population section. The denominator for this calculation will be women age 11 to 55 in the database.

### 4.3.2 Time to Discontinuation

The study is in estimating the duration (in days) that a woman is continuously prescribed a particular treatment (or conversely the time until they stop the treatment). This duration is determined by the length of time for a particular treatment era. The term era is used to refer to the span of time a person is considered “exposed” to a treatment of interest. Discontinuation of a treatment is identified as a switch to a different drug or the woman has turned 55 years old. Any gap more than **30** days would be considered treatment discontinuation. An observation is censored if a switch is not determined based on end of continuous observation or end of the study period.

# 5. Data Sources

The datasource 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** **:** Data Sources used in study

| Name | Country | Type | Size | Availability | Description |
| --- | --- | --- | --- | --- | --- |
| CPRD AURM | UK | EHR | Number | 2008-2010 | Primary care data from a participating electronic system in England |
| CPRD AURM | UK | EHR | Number | 2008-2010 | Primary care data from across the UK |

**?(caption)**

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 a observation window of 365 to 0 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. Table xxx provides the set of covariates to be assessed in the study.

**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) in following categories
  + Underweight ()
  + Normal Weight ()
  + Overweight ()
  + Obese ()
  + Morbidly obese ()
* Charlston Comorbidity Score
* CHADs2Vasc

**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. The roll-up is used for the table 1 report and the individual concepts are used to construct manhattan plots.

**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
  + Hysterectomy
  + Endometrial ablation
  + Uterine artery embolization (UAE)
  + Myomectomy
  + Blood transfusion

## 6.3 Post-Index Utilization

We assess the point prevalence of drug use at time intervals of 0 days, 1 to 183 days, 184 to 365 days, 366 to 730 days, and 731 to 1825 days post index. Point prevalence is defined as the number of affected persons present in the population at a certain time divided by the number of persons in the population at that time. For the post-index utilization we focus on drugs and procedures, listed below.

* Drugs
  + 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 [Ulipristal acetate (G03AD02, G03XB02)]
  + Danazol (G03XA01)
  + Gonadotropin releasing hormone analogues (L02AE)
  + Intrauterine devices (G02BA)
  + Iron preparations (B03A)
* Procedures
  + Hysterectomy
  + Endometrial ablation
  + Uterine artery embolization (UAE)
  + Myomectomy

## 6.4 Treatment Patterns

For the primary objective, we are interested in assessing treatment sequences of pharmacotherapy and surgery used to treat patients with HMB. Enumeration of the sequences will highlight distinct treatment pathways used by women in the target cohort. The list of treatment patterns is the same as that of the post-index utilization section. Construction of the treatment patterns analysis is based on the framework detailed by Markus et al (Markus et al. 2022). A treatment pathway is the sequence of treatments starting from the index date and ending at the end of the study period. When defining treatment pathways, the term era is used to refer to the span of time a person is considered “exposed” to a treatment of interest. An era is constructed based on rules that combine successive periods of drug exposure into a continuous period. The minimum era duration determines the minimum duration for an era to be considered in the study; for this study there is no minimum era duration. The maximum time interval between two subsequent eras of the same treatments to collapse into the same era is **30** days. This means that if there is a gap greater than **30** days between the same treatment, this would indicate two separate eras in the pathway. A treatment combination is defined as an interval when a patient has two treatment eras overlapping in time. The minimum overlap time required between multiple treatments to be considered a combination is **30** days. The minimum time that a treatment era (before or after a combination of treatments) must last in order to be included as a separate treatment in the pathways is **30** days. This means that if treatment is observed for less than **30** days before or after a combination it is not considered as part of the sequence. Finally, only treatment changes are reported in the path, meaning we remove repeated treatments in the pathway and only report when a change occurs. The treatment sequence results will be displayed using a sankey diagram, complemented by a summary table that shows the enumeration of each specific sequence. Sequences will be dropped from the analysis if the count is less than **20**.

## 6.5 Time to Event

Using the outcome definition for time to discontinuation we will estimate the Kaplan-Meier (KM) curve for the treatment events specified in the post-index utilization section. From the KM we report the probability of survival at 183-, 365- and 730- days as well as the median, 25th and 75th percentile time to event. All of these statistics are reported with a corresponding two-sided 95% confidence interval.

## 6.6 Incidence Analysis

In this study we will calculate the incidence of the HMB cohort in a population of women age 11 to 55. 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.

## 6.7 Stratafication

No stratifications will be made in this analysis.