SAP for EHDEN HMB

Martin Lavallee

Siir Su Saydam

Asieh Golozar

2023-08-04

Table of contents

[1. Background 1](#_heading=h.30j0zll)

[2. Research Questions 2](#_heading=h.3znysh7)

[3. Objectives 2](#_heading=h.tyjcwt)

[3.1 Primary Objectives 2](#_heading=h.1t3h5sf)

[4. Study Design 2](#_heading=h.2s8eyo1)

[4.1 Study Population 3](#_heading=h.3rdcrjn)

[4.2 Exposure Definition 3](#_heading=h.lnxbz9)

[4.3 Outcome Definition 3](#_heading=h.1ksv4uv)

[5. Data Sources 3](#_heading=h.2jxsxqh)

[6. Analysis Plan 4](#_heading=h.1y810tw)

[6.1 Cohort Diagnostics 4](#_heading=h.2xcytpi)

[6.2 Baseline Characteristics 4](#_heading=h.3whwml4)

[6.3 Treatment Landscape 5](#_heading=h.qsh70q)

[6.3.1 Prevalence of Drug Exposures 6](#_heading=h.1pxezwc)

[6.3.2 Constructing the treatment history 7](#_heading=h.2p2csry)

[6.3.3 Treatment Pathways 7](#_heading=h.3o7alnk)

[6.3.4 Duration of Drug use 7](#_heading=h.ihv636)

[6.4 Procedure use post-index 7](#_heading=h.1hmsyys)

[6.5 Incidence Analysis 8](#_heading=h.2grqrue)

# 1. Background

Heavy menstrual bleeding (HMB) is a symptom that has been defined as an “excessive menstrual blood loss, which interferes with a woman’s physical, social, emotional and/or material quality of life, and which can occur alone or in combination with other symptoms” [1–4]. Large real-world population-based studies are not currently available to quantify and comprehensively characterize women with HMB with respect to age at diagnosis, reproductive history and reproductive age, underlying conditions associated with HMB, as well as other comorbidities and comedications.

The concept for this study originated from a study proposal awarded by the European Health Data and Evidence Network (EHDEN), a public-private consortium of the Innovative Medicine Initiative (IMI). EHDEN is creating a federated network of real-world data (RWD) sources across Europe, all of which are standardized to the Observational Medical Outcomes Partnerships (OMOP) common data model (CDM), which maps diverse database codes to a standardized set of concepts. The use of standardized OMOP CDM data is particularly relevant for this study, as it enables the characterization of HMB patients and their treatment pathways across multiple countries and healthcare systems.

This study is designed in two parts. Part 1 of this study is conducted in internally by bayer for describe incidence of HMB, clinical characteristics of women diagnosed with HMB and treatment pathways in five healthcare databases covering the UK, Germany, France and the US. This study is later amended for the currently proposed study (Part 2) to be released for execution in additional countries within EHDEN. The study will generate evidence to identify women diagnosed with HMB in respective databases, as well as these patients’ baseline characteristics and longitudinal treatment pathways across countries with different healthcare systems. The purpose of this document is elaborate on the study design constructed by Odysseus Data Services. A separate protocol has been written by Bayer.

# 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 of women diagnosed with 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.

## 3.2 Exploratory Objectives

The exploratory objectives of this study are to:

* Estimate the incidence of HMB per calendar year and over the entire study period in selected data sources.
* Describe country-specific guideline compliant treatment use where national treatment guidelines are available for HMB.

# 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 assets (CPRD GOLD, CPRD Aurum, Optum Claims, MarketScan) 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 ends 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, therefore 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 of heavy-menstrual bleeding (index event) on or after January 1, 2000 and up to the date of latest data availability at the time of study conduct. Study end date might show variability across data sources. 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:

* Death
* Turning age 55 (indicating transition into natural menopause)
* Observation of a hysterectomy or bilateral ovariectomy
* Observation of a menopause diagnosis

### 4.1.1 Incidence Denominator Cohort

For the incidence calculation, the denominator is women with any kind of healthcare (index event) observed between the years of 2000 and 2022. Women in the denominator cohort must have a minimum of 365 days of prior observation.

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

* end of continuous observation in the database
* any kind of healthcare interaction at age > 55 years
* menopause (induced or natural)
* hysterectomy or bilateral oophorectomy
* uterine or ovarian cancer
* pregnancy
* death

Concepts and logic for the cohort definitions of the study population and incidence denominator will be supplied in the *CohortDetails* supplementary file.

## 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 data sources for this study include a) Bayer OMOP assets of CPRD Aurum, CPRD GOLD, Optum Claims, MarketScan 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  Aurum | UK | EHR |  |  | Primary care data from a participating electronic system in England |
| CPRD  GOLD | UK | EHR |  |  | Primary care data from across the UK |
| Optum Claims | US | Claims |  |  | US commercial claims covering all 50 states |
| MarketScan | US | Claims |  |  |  |
| Hospital del Mar | Spain | EHR |  |  | Hospital data from Barcelona Spain |
| Ministry of Health | Israel | EHR |  |  |  |
| THIN | Belgium | EHR |  |  |  |
|  | Germany | EHR |  |  |  |
|  | Finland | EHR |  |  |  |
|  | Hungary | EHR |  |  |  |
|  | Norway | EHR |  |  |  |

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

# 6. Data Analysis

## 6.1 Cohort Diagnostics

Prior to running any specific analysis, we will evaluate the HMB cohort using the OHDSI R package CohortDiagnostics[5]. 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 0 days prior to the index date, unless specified otherwise. Categorical covariates are reported as counts and percentages. Continuous covariates are reported as median, interquartile range (25th and 75th percentile), minimum and maximum.

**Demographics and clinical characteristics**

* Age at HMB diagnosis as 5-year categories
* Age at HMB diagnosis as continuous variable
* Race (if available)
* Ethnicity (if available)
* Year of HMB Diagnosis (per calendar year)
* Comorbidity scores
  + Charlson Comorbidity Index (CCI) (continuous)
  + CHA2DS2-VASc score (continuous)
* **BMI (if available) continuous and in following categories within 365 days before and including index date[[1]](#footnote-1)**
* 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)
* Missing

**Concept-based**

For analysis on data converted to the OMOP CDM, covariates can be assessed based on either singular standardized concepts from the OMOP Vocabulary or by bundling multiple concepts through cohort definitions. A concept-based covariate describes the occurrence of an event based on the presence of a standardized code in the OMOP CDM. For conditions, standardized concepts are based on Systematized Nomenclature of Medicine Vocabulary (SNOMED) and its hierarchy; including descendant concepts that are linked to the top-level term. For drugs, standardized concepts are based on the RxNorm vocabulary, where the hierarchy is rolled up to classes of the Anatomical Therapeutic Chemical (ATC) vocabulary. Assessing concept-based covariates, in an OMOP analysis, is a general approach to assessing the clinical features of a population because it considers a very large set of features with minimal user specification.

* 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[6].

**Cohort-based**

Cohort-based covariates, in an OMOP analysis, are more specific definitions of clinical concepts of interest that are built using multiple concepts from the OMOP standardized vocabulary. Cohort-based covariates are used in the context of this study to specifically identify conditions of interest, determined prior to the study execution, that are not sufficiently identified via single concepts in the OMOP vocabulary.

* Comorbidities
  + Diabetes Mellitus
  + Polycystic Ovary Syndrome (PCOS)
  + Dysmenorrhea
  + Pain
  + Anemia
  + Iron deficiency anemia
  + Obesity
* Comedications
  + Antithrombotic agents
  + Antidepressants
  + Tamoxifen
  + Antipsychotics
  + Gonadal steroids
  + Tranexamic acid
  + Progestin only regimens
    - Medroxyprogesterone acetate (MPA)
    - Oral norethindrone acetate (NETA)
    - Desogestrel
    - Etonogestrel implant
    - Drospirenone
  + Non-steroidal anti-inflammatory drugs (NSAIDs)
  + Combined oral hormonal contraceptives
  + Ulipristal acetate
  + Danazol
  + Gonadotropin releasing hormone analogues
  + Intrauterine devices (IUD)
    - Plastic IUD with progestogen
    - Plastic IUD with copper
  + Iron preparations
* Procedures
  + Endometrial ablation
  + Uterine artery embolization (UAE)
  + Myomectomy
  + Blood transfusion
  + Intrauterine devices (recorded as procedures)
    - Hormonal IUD
    - Copper IUD
    - Undefined IUD

Concepts and logic for these cohort definitions will be supplied in the *CohortDetails* supplementary file.

**Underlying Causes of HMB**

In addition to assessing characteristics at baseline, proportion of patients with underlying causes of HMB will be calculated. Conditions that are considered to be underlying causes of HMB will be characterized prior 365 days prior to and including index as well as post-index in the following time intervals: 1 to 365 days, 366 to 730 days, and 731 to 1825 days from index date. Reported counts will be limited to the first occurrence of the underlying cause. The underlying causes of HMB are:

* Uterine endometrial polyps
* Adenomyosis
* Uterine leiomyoma
* Endometrial hyperplasia
* Coagulopathy or coagulation disorders
* Ovulatory dysfunction (including hypothyroidism, polycystic ovary syndrome, adrenal disorders, hyperprolactinemia, hypothalamic disorders)
* Endometriosis

The logic and concept set for the cohort definitions for these underlying causes of HMB will be supplied in the *CohortDetails* supplementary document.

## 6.3 Treatment Patterns

In this study we assess the treatment patterns for female patients with HMB which encompasses three metrics: first, the prevalence of drug use at pre-defined time windows post-index; second, the patterns of treatments including single drugs and combinations of treatment (later referred to as treatment pathways); and third, the length of time spent taking the drug exposures of interest, both as single exposures or in combination (time to treatment discontinuation). The drug exposures of interest are:

* Tranexamic acid
* Progestin only regimens:
  + Medroxyprogesterone acetate (MPA)
  + Oral norethindrone acetate (NETA)
  + Desogestrel
  + Etonogestrel implant
  + Drospirenone
* Non-steroidal anti-inflammatory drugs (NSAIDs)
* Combined oral hormonal contraceptives
* Ulipristal acetate
* Danazol
* Gonadotropin releasing hormone analogues
* Intrauterine devices (Plastic IUD with progestogen)

For the treatment patterns analysis, a set of drug exposure cohorts will be defined, where the index event of each cohort is prescription of one of the drugs of interest during the study period. People remain in the cohort if they satisfy the selection criteria for the study cohort. A person's 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. A maximum of 30 days between drug exposure records are allowed to consider these exposures as a 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:

* end of continuous observation in the database
* any kind of healthcare interaction at age > 55 years
* menopause (induced or natural)
* hysterectomy or bilateral oophorectomy
* uterine or ovarian cancer
* pregnancy
* death

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.

### 6.3.1 Prevalence of Drug Exposures

The first metric under treatment patterns 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 the prevalence of drug exposures for women in the HMB cohort are: at index, 1 to 183 days, 184 to 365 days, 366 to 730 days, and 731 to 1825 days post index. Counts and percentages for each drug exposure cohort (from the listed above) will be reported at each time window.

### 6.3.2 Treatment Pathways

For each female with HMB, the sequence of treatments taken throughout the patient history will be characterized. The first 15 sequences will be reported for each database where each sequence must have a minimum of 10 unique persons to be included in the treatment pathways report. The treatment sequence results will be displayed using a Sankey diagram, complemented by a summary table that shows the enumeration of each specific sequence.

### 6.3.3 Time to Discontinuation

Using the treatment history, we calculate the median (and 95% confidence interval) time spent on a specific drug using the Kaplan-Meier (KM) methodology. Events are censored if the persons had any kind of healthcare interaction at age > 55 years, menopause (induced or natural), hysterectomy or bilateral oophorectomy, uterine or ovarian cancer, pregnancy or all-cause mortality. Combinations of treatments will be handled as its own era in the calculation of time to discontinuation.

## 6.4 Treatment Procedures

In this analysis, procedures used to treat HMB will also be characterized post-index during the study period. 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.

We report the prevalence of each procedure during the following time intervals: 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 interval. In addition we report the median, 25th and 75th percentile of the duration of time to intervention.

## 6.5 Exploratory Analyses

## 6.5.1 Incidence Analysis

The overall and yearly incidence rate of HMB is calculated in each participating database. The numerator in the incidence calculation is the HMB target cohort defined in Section 4.1 and the denominator is the cohort of women with a healthcare visit as defined in Section 4.1.1. Incidence rate is defined based on the following formula:

The incidence will be calculated using the IncidencePrevalence package from DARWIN (3).

## 6.5.2 Guideline Compliance

# Bibliography

1. Munro MG, Critchley HOD, Fraser IS, FIGO Menstrual Disorders Committee. The two FIGO systems for normal and abnormal uterine bleeding symptoms and classification of causes of abnormal uterine bleeding in the reproductive years: 2018 revisions. Int J Gynaecol Obstet Off Organ Int Fed Gynaecol Obstet 2018;143:393–408. <https://doi.org/10.1002/ijgo.12666>.
2. Karlsson TS, Marions LB, Edlund MG. Heavy menstrual bleeding significantly affects quality of life. Acta Obstet Gynecol Scand 2014;93:52–7. <https://doi.org/10.1111/aogs.12292>.
3. Weyand AC, Fitzgerald KD, McGrath M, Gupta V, Braun TM, Quint EH, et al. Depression in Female Adolescents with Heavy Menstrual Bleeding. J Pediatr 2021;0. <https://doi.org/10.1016/j.jpeds.2021.09.007>.
4. Shimamoto K, Hirano M, Wada-Hiraike O, Goto R, Osuga Y. Examining the association between menstrual symptoms and health-related quality of life among working women in Japan using the EQ-5D. BMC Womens Health 2021;21:325. <https://doi.org/10.1186/s12905-021-01462-7>.
5. Rao G, Schuemie M, Ryan P, Weaver J, Gilbert J. CohortDiagnostics: Diagnostics for OHDSI Cohorts. 2022.
6. Schuemie M, Suchard M, Ryan P, Reps J, Sena A. FeatureExtraction: Generating Features for a Cohort [Internet]. 2021. Available from: https://github.com/OHDSI/FeatureExtraction
7. Edward Burn, Berta Raventos and Marti Catala (2023). IncidencePrevalence: Estimate Incidence and Prevalence using the OMOP Common Data Model. R package version 0.4.1. https://darwin-eu.github.io/IncidencePrevalence/

1. If more than one BMI value is available, the closest occurrence to the index date will be reported. [↑](#footnote-ref-1)