EPIDEMIO-STRATEGY & MEDICAL ECONOMICS (ESME) Research Program

Academic Real World Data Platform

Metastatic Breast Cancer

Reference of ISO document	N/A
Plan	PGP
Project	ESME MBC

Document title	Methodology
Date	07-DEC-2023
Version	1.6

File name	ESME CSM – PGP – A16 – Methodology_v1.6_2023.12.07

Table of contents

1 Introduction

The Epidemio-Strategy and Medical Economic (ESME) Research Program has been initiated and is managed in compliance with the recommendations/guidelines Good Epidemiology Practices (GEP) and Good Pharmacoepidemiology Practices (GPP). Unicancer manages this program which centralizes data on patients suffering from metastatatic breast cancer (MBC) treated at one of the 18 French Comprehensive Cancer Centres (FCCC).

This document details the methodology, data quality, project management and legal provisions.

2 Project Methodology

2.1 Source population: selection criteria

Cases are selected to enter the ESME MBC database according to specific criteria:

- Male or female
- Aged ≥18 years
- Patient with metastatic breast cancer whose first metastasis was treated (either completely or partially) in a FCCC from January 1, 2008. Treatment strategies could include radiotherapy, chemotherapy, targeted therapy or endocrine therapy.

2.2 Patient selection process

The rigorous process of patient selection involves both screening and selection phases, and is based on information retrieved from multiple data sources within each FCCC. The objective of the screening phase is to identify all patients potentially meeting the database selection criteria, including patients who are treated without hospitalisation or surgery (i.e. patients not listed in structured databases within the FCCCs).

The overall aim, therefore, is to create a list of screened patients that is as comprehensive as possible while minimising the number of false positives. In each FCCC, patients are identified from hospitalisation records. Depending on the software used in the centre's pharmacy information system, patients may also be identified from pharmacy records. Other data sources, such as MBC-specific or cancer databases and multidisciplinary team meeting

records are additionally used for patient screening, as are search engines in the patient electronic file using relevant keywords.

Structured data source used to generate the screening list are the following:

- French medical information system (called PMSI)
- Pharmacy records
- Other databases or serach engines

2.2.1 Data source from French medical information system

The French medical information system (called PMSI) is a mandatory data structure describing each inpatient stays and homogenous across all health care establishments in France in order to invoice the National Health Insurance Fund. The hospitalisation records are structured using the MBC-specific international classification of diseases (ICD) codes associated with inpatient stays.

All variables related to patients and their stay allow a ranking (using modelling) of the hospital stay and are associated to lump sum category.

Each hospital stay is data-structured called Standardized Hospital Discharge Summary. Each Standardized Hospital Discharge Summary is composed of one or several medical unit summaries. Data structuration is defined as subsets following a standardized process (formats, dictionaries, etc) and main subsets are the following:

- Medico-administrative variables
- Diagnoses
- Medical procedures
- Other specific variables

The subset « Diagnoses » is composed of clinical diagnostic elements described on a hierarchical manner with an attribute. Each diagnosis/attribute couple is defined as follows:

- Main diagnosis
- Linked diagnosis
- Related significant diagnoses or documentary related diagnoses

Coding of diagnostic elements is performed using the World Health Organisation (WHO) ICD.

Coded diagnoses for each hospital stay for each centre are used as data source. The following diagnostic codes are used to identify cases:

Diagnostic code	ICD10 Code Title
C50.-	Malignant neoplasm of breast
C77.- (except C77.3)	Secondary and unspecified malignant neoplasm of lymph nodes
C78.-	Secondary malignant neoplasm of respiratory and digestive organs
C79.-	Secondary malignant neoplasm of other and unspecified sites

2.2.2 Data source from pharmacy records

In all FCCC, computerized pharmacy database contains records on all dispensed medications by the medical centre. According to the software setttings, some of the pharmacy databases may be used for patient screening.

Screening is performed using the medication dispensations over the selection period and related to one or more criteria as follows:

- A diagnostic related to MBC (variable or code)
- A specific information (metastatic treatment line or palliative staging for a BC)
- As a default, chemotherapy protocols specific for MBC care (sensitivity and specificity are centre-specific)

2.2.3 Other data sources: other databases or search engines

The screening may use other tools such as:

- Local BC-specific database or local cancer database
- Engine search tools using key words (breast + metastasis + endocrine therapy + oral therapy)
- Structured database containing records on multidisciplinary meetings specific to MBC

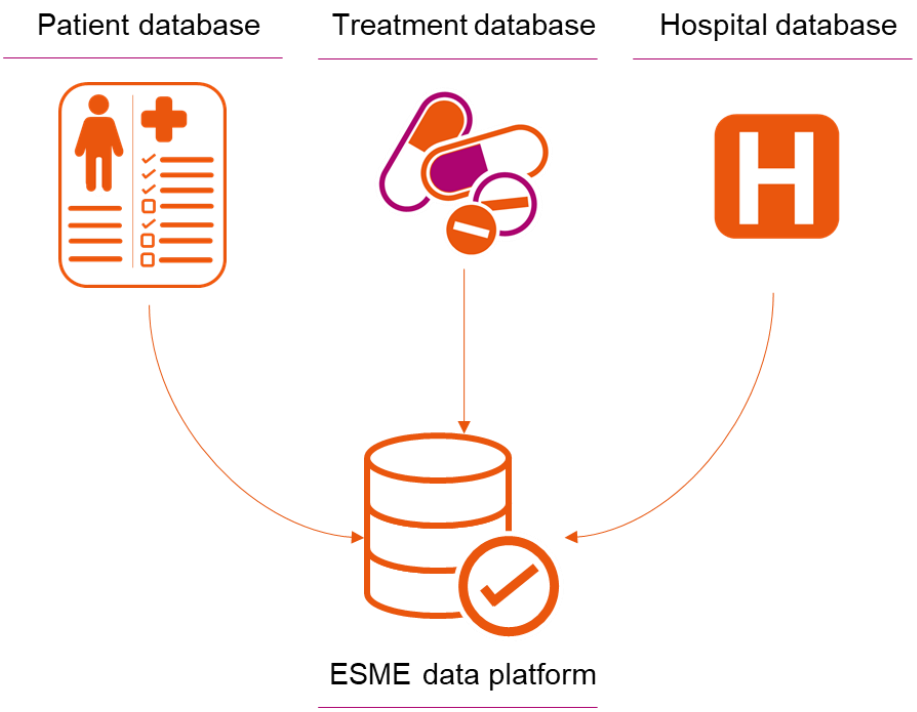
2.3 Patient selection process

Once the screening list is finalised at each centre, each identified case is given an anonymous ESME number. Electronic medical records (EMR) for each screened patient is reviewed to check the selection criteria aforementioned (see section 2.1).

2.4 Specifications of the data platform

The ESME MBC data plateform compiles data from patient's Electronic medical records, inpatient hospitalisation records and pharmacy records (Figure 1).

Figure 1. Data sources used to constitute the ESME MBC database



Patient-related data integrates data from non-structured sources, such as patient medical records and reports of multidiscipline team meetings. It provides information on patient demographics, primitive tumour, relapses, histology, metastatic disease, therapeutic care/settings, withdrawal reasons, and clinical events. Specifications (list of variables) are available as an appendix in the Data Management Plan (DMP).

Pharmacy records-related data is a structured database including all data pertaining to medications prescribed and dispensed by pharmacies within each FCCC, specifically

chemotherapy and co-prescriptions (dates, specialties, treatment protocols, etc.). It does not contain information on products that are prescribed or delivered outside of the centre. Specifications (list of variables) are available as an appendix in the DMP.

Hospitalisation-related data is a structured and systematic database that contains all medical data related to any hospitalisation, and is used to bill the French National Health Insurance Fund (Assurance Maladie). It provides information on patient entry and discharge dates as well as diagnostic and therapeutic procedures performed during patient stay, including radiotherapy and surgery. Specifications (list of variables) are available as an appendix in the DMP.

3 Data Quality

3.1 Project management process

Unicancer defined quality standard for project management as follows:

- Data collection on exhaustive population and data available according to the list of variables defined for the data platform;
- Compliance with recommendations for best practices applicable to the fields;
- Data control, validation and coding of centralized data;
- Unique electronic tool for data collection with dynamic edit checks;
- Supervision of expert groups for the planning, execution and follow-up of the project;
- Random on-site quality review process on all screened cases;
- Annual (internal and external) audit programme: data management, project management, on-site operations, project providers

3.2 Data management

Any data integrated in the ESME Resarch Program follows quality-controlled processes. Patients-related data is collected in each centre by technicians who are specifically trained for the project using an electronic data collection (eDC) tool. Medical support is provided for training and assisting data collection staff, and an expert panel is available to answer any questions regarding patients' records. Data imported into the final database is controlled, recoded, and harmonized according to a Data Management Plan. There is no transmission of

individual data; data is centralized within each centre using a shared anonymous format. All data is exclusively obtained retrospectively. No attempt is made to recover missing/not available data from the patients' medical record by contacting healthcare providers or patients.

The ESME MBC database is locked for each annual update and made available for analyses on a yearly basis. Databases and the eDC tool are maintained by a certified host of health databases. Unicancer coordinates secure access to the global ESME Data Platform and is.

3.3 On-site quality review

Two types of reviews are performed in each FCCC. These reviews consist in double checking the validation of selection criteria for selected and non-selected cases.

For non-selected cases, selection criteria and reason for non-selection are double-checked *versus* the source data. Accepted error limit is 10%.

For selected cases, selection criteria and key variables entered in the database are also cross-checked *versus* the source data. Accepted error limit is 5%.

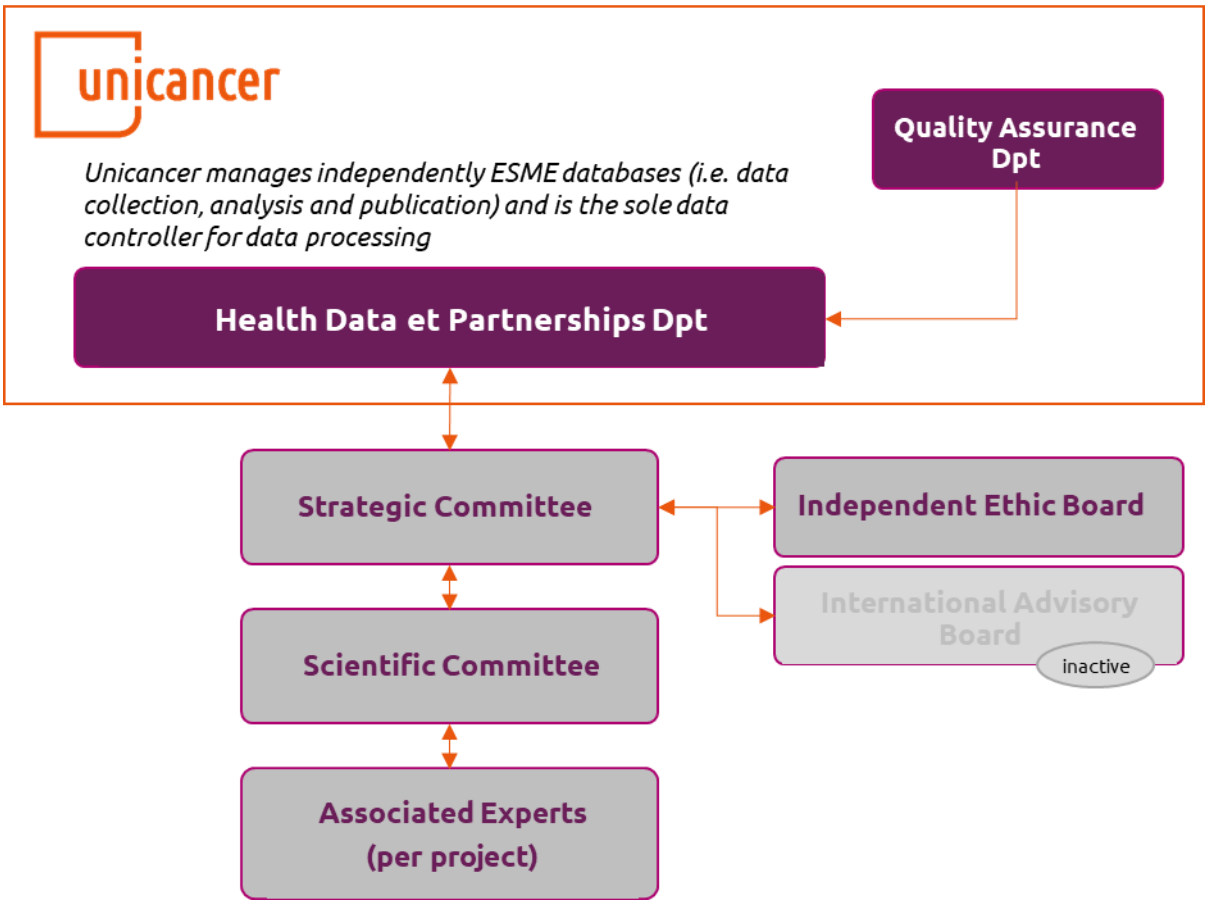
3.4 Audits

Unicancer Quality Assurance Department conducted a central audit at local level (medical centres) on data entry in the patient database and on the constitution of preselection lists, as well as on service provider (e.g. statistics activities). A yearly audit plan systematically includes centre's audits and external service provider audit.

4 Governance

Three Committees have been established in the ESME Research Program: the ESME Strategic Committee, the ESME Scientific Committee and the ESME independent ethic board. Responsibilities and tasks of each committee are listed below in Sections 4.1, 4.2 and 4.3. Tasks and operating rules of each Committee are defined in specific charters.

Figure 2. ESME Governance



4.1 The ESME Strategic Committee

- Defines the ethical and scientific charter of the program
- Defines a consistent strategy between the actors involved (participating centers, Unicancer, public and private partners)

- Suggests initiatives and collaboration perspectives for the Program
- Suggests technical solutions to enable the implementation of long-term operations within the platforms (data collection, updates, analyses, etc.)

4.2 The ESME MBC Scientific Committee

- Ensure the application of the scientific strategy defined for the platform
- Evaluate requests for project research nto be conducted on the platform in compliance with the eligibility criteria defined in the program's operating charter
- Provide recommendations on analysis methods
- Guarantee the scientific relevance of the results published via the review of all communications made using data from the ESME MBC platform

4.3 The ESME Independent Ethic Board

- Manages conflicts of interest
- Provide recommendations that may improve the system in general and prevent risks related to conflicts of interest
- Gives an opinion on individual or specific situations
- Gives an opinion on possible collaborations with private partners

5 Project Management

The ESME MBC data platform project is managed by a dedicated team at Unicancer and a local coordinator at central and local levels respectively. The local coordinator is appointed by the general manager (GM) of each FCCC.

Retrospective data collection contributing to the ESME MBC data platform is under GM responsibility and no health care practitioner is reached out to provide EMR with additional data that is not initially documented.

The local coordinator, main contact for Unicancer, is the person in charge of coordinating all actions necessary to ensure a smoothy project management at a local level:

- Management of the screening lists (data sources identifications, anonymisation and maintenance) for each selection period
- Administrative and technical management of all resources involved in the patient selection process and data collection for patient database
- Adequate data transfer for treatment and hospitalisation databases

6 Legal Provisions

The applicable legal provisions are observed. In compliance with French regulations, the ESME MBC data plateform was authorized by the French data protection authority (authorization no. 1704113 and subsequent amendments). In compliance with the applicable European regulations, a complementary authorization was obtained on October, 2019.

Unicancer manages the ESME Data Warehouse (including the ESME MBC data platform) in accordance with current best practice guidelines.