NON-INTERVENTIONAL STUDY PROTOCOL

Study Title:	**Real-world Study on Bemiparin Effect in Patients with Cancer-Associated Thromboembolism Using Artificial Intelligence (BEMICAT Study)**
Study code	ROV-BEMI-2024-01
Study Sponsor:	LABORATORIOS FARMACÉUTICOS ROVI, S.A. C/Alfonso Gómez, 45 Madrid 28037 Spain
Version:	1.2
Date:	11JUN2025

Approval Section: *Signatures boxes should be fulfilled including name, role, and date in addition to the signature.*

Javier Martínez González ROVI Medical Director Date: 17-Jun-2025	Ana Belén López Yélamos ROVI EU-QPPV Date: 17-Jun-2025
Carlos Arias Medical Research Expert Project Lead (Savana Research, S.L.) 17-Jun-2025	David Casadevall Head of NLP Analytics and Biostatistics Medsavana, S.L. on behalf of Savana Research, S.L.) 17-Jun-2025
Natalia Polo Project Manager (Medsavana, S.L. on behalf of Savana Research, S.L.) 17-Jun-2025	Eduard Sarró Medical Documentation Expert Project Lead (Savana Research, S.L.) 17-Jun-2025

CONFIDENTIAL

This protocol is the confidential information of Laboratorios Farmacéuticos ROVI, S.A. and is intended solely for the guidance of the clinical investigation. This protocol may not be disclosed to parties not associated with the clinical investigation or used for any purpose without the prior written consent of Laboratorios Farmacéuticos ROVI, S.A.

CONFIDENTIALITY STATEMENT

This protocol contains information that is confidential and proprietary to Laboratorios Farmacéuticos ROVI, S.A. ("Confidential Information"). This Confidential Information is being provided to you for the purpose of conducting a clinical study for Laboratorios Farmacéuticos ROVI, S.A. You may only disclose the contents of this protocol to study personnel under your supervision who need to know the contents for this purpose, as well as to your Institutional Review Board(s) or Ethics Committee(s)("Authorised Persons"), subject to the following conditions: (i) You must inform such Authorised Persons about the confidential nature of the Confidential Information and ensure they all personally assume a likewise obligation of confidentiality; (ii) You will be liable to Laboratorios Farmacéuticos ROVI, S.A. for any breach of this confidentiality obligation by such Authorised Persons; (iii) The contents of this protocol may not be disclosed to any person or entity without the prior written permission of Laboratorios Farmacéuticos ROVI, S.A. and may not be used for any purpose other than the conduct of this study. The foregoing shall not apply to disclosure required by governmental regulations or laws; however, you will give prompt notice to Laboratorios Farmacéuticos ROVI, S.A. of any such disclosure and permit Laboratorios Farmacéuticos ROVI, S.A. to take steps to avoid the need for such disclosure and/or to obtain a protective order governing same.

Any supplemental information that may be added to this document is also confidential and proprietary to Laboratorios Farmacéuticos ROVI, S.A. and must be kept in confidence in the same manner as the contents of this protocol.

Any person who receives this protocol without due authorization from Laboratorios Farmacéuticos ROVI, S.A. is requested to return it to Laboratorios Farmacéuticos ROVI, S.A. or to promptly destroy it and inform Laboratorios Farmacéuticos ROVI, S.A. upon destruction.

Laboratorios Farmacéuticos ROVI, S.A. makes no representations or warranties of any kind, express or implied, concerning the Confidential Information, and the absence of any actual or legal defects, whether or not discoverable. Specifically, and not to limit the foregoing, Laboratorios Farmacéuticos ROVI, S.A. makes no warranty or representations:

i. regarding the correctness, completeness or fitness for a particular purpose of the Confidential Information, and

ii. that the use of the Confidential Information will not infringe any patents or other intellectual property rights or copyrights of a third party, and

iii. that the use of the Confidential Information will not cause any damage of any kind to the Receiving Party or a third party.

Confidential Information is disclosed by Laboratorios Farmacéuticos ROVI, S.A. on a non-exclusive basis and it remains free to use and disclose Confidential Information with third parties at any time at its own option.

You will notify Laboratorios Farmacéuticos ROVI, S.A. immediately upon discovery of any unauthorized use or disclosure of any piece of Confidential Information or any other breach of this covenant by you. You will cooperate with Laboratorios Farmacéuticos ROVI, S.A. in every reasonable way to help it regain possession of the leaked Confidential Information and prevent its further unauthorized use.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTAL TABLES

VERSION HISTORY

VERSION	DATE	CHANGE(S) SINCE LAST VERSION
1.0	18DEC2024	First version
1.1	27FEB2025	Adapted version
1.2	11JUN2025	Typo corrections and addition of complementary references

1 LIST OF ABBREVIATIONS

Table 1: List of abbreviations and definitions of terms

Abbreviation or term	Definition
AE	Adverse Event
AI	Artificial Intelligence
CAT	Cancer-associated thromboembolism
CRNMB	Clinically relevant non-major bleeding
DOACs	Direct oral anticoagulants
DVT	Deep vein thrombosis
EHR	Electronic Health Records
IRB	Institutional Review Board
IEC	Independent Ethics Committee
LMWH	Low-molecular-weight heparin
ML	Machine Learning
NEL	Name Entity Linking
NER	Name Entity Recognition
NLP	Natural Language Processing
RWD	Real-World Data
RWE	Real-World Evidence
PE	Pulmonary embolism
PO	Primary Objective
SAP	Statistical Analysis Plan
SO	Secondary Objective
TFLs	Tables, Figures & Listings
VKA	Vitamin K antagonists
VTE	Venous thromboembolism

2 RESPONSIBLE PARTIES

This study is sponsored by Laboratorios Farmacéuticos ROVI S.A., located at C/Julián Camarillo, 35, 28037 Madrid, Spain (hereinafter referred to as the "Sponsor").

The Sponsor has contracted Medsavana S.L. (hereinafter "Medsavana"), located at Calle Larra, 12, 28004, Madrid, Spain, along with its affiliate Savana Research S.L. (hereinafter "Savana Research"), operating in the same fiscal address, to conduct the study and analyze the results.

A complete list of principal investigators and sites involved in the study will be maintained in a separate document.

3 STUDY SUMMARY

Title	Real-world Study on Bemiparin Effect in Patients with Cancer-Associated Thromboembolism Using Artificial Intelligence (BEMICAT Study)
Phase	**IV (Non-Interventional Study)**
Type	**Retrospective**
Keywords	Cancer-associated thromboembolism; deep vein thrombosis, pulmonary embolism; Bemiparin electronic health records; natural language processing; artificial intelligence.
Rationale and background	Cancer-associated thromboembolism (CAT) is a major health concern and a leading cause of death in cancer patients, despite being preventable (1, 2). CAT typically manifests as venous thromboembolism (VTE), including deep vein thrombosis (DVT) and pulmonary embolism (PE), and is linked to the hypercoagulable state caused by cancer (3). Cancer patients have a significantly higher risk of VTE, with its incidence tripling in the past 20 years (1, 4). Reported CAT rates range from 2% to 30%, depending on tumor type, location, metastases (5-7), treatments (8-10), and patient factors (7). Patients with CAT have a two- to threefold higher mortality rate (4). VTE recurrence within six months is 5.6% to 7.9%, with an overall prevalence of 13 cases per 1,000 person-years (11). VTE recurrence in cancer patients is associated with increased morbidity and higher healthcare use (12). Current guidelines recommend long-term anticoagulation therapy with low-molecular weight heparin (LMWH) or direct oral anticoagulants (DOACs) for at least six months (1, 13). However, cancer patients on anticoagulation face a higher risk of major bleeding, requiring individualized assessments of VTE recurrence and bleeding risks before starting treatment (14, 15). LMWHs remain the preferred therapy in patients with CAT in real life. A study using the TESEO registry showed 88% of CAT cases were treated with any of the available LMWHs being Enoxaparin the most used (16). Although Bemiparin has shown effectiveness and safety in clinical practice (17-19), no clinical trials directly compare its outcomes with other LMWHs like Tinzaparin or Enoxaparin in this population. Medsavana's EHRead® technology (12, 20-27), which utilizes natural language processing (NLP) and machine learning (ML), provides an opportunity to analyze the effectiveness and safety of Bemiparin compared to other LMWHs in CAT management.
Objectives	***Phase I*** To assess in a limited sample whether the incidence of CAT in patients who started long-term treatment with Bemiparin, Dalteparin, Enoxaparin or Tinzaparin during the recruitment period is consistent with the sample size estimated for phase II. ***Phase II*** **Primary Objectives (PO)** ▪ **PO1**: To assess the noninferiority of long-term Bemiparin treatment in preventing VTE recurrences in adult patients with CAT over a six-month period after the index CAT episode, compared to pooled data of Dalteparin, Enoxaparin, and Tinzaparin. ▪ **PO2:** To compare the safety of long-term Bemiparin treatment versus pooled data of Dalteparin, Enoxaparin, and Tinzaparin in adult patients with CAT over a six-month period after the index CAT episode. **Secondary Objectives (SO)** In patients with CAT on long-term anticoagulation with the drugs Bemiparin, Dalteparin, Enoxaparin, or Tinzaparin (overall and stratified by drug type): ▪ **SO1:** To describe the sociodemographic and clinical characteristics, including known risk factors present at the index CAT episode. ▪ **SO2:** To describe the patient journey and treatment pattern, including anticoagulation switches. ▪ **SO3:** To describe the time-to-event for effectiveness and safety outcomes.

		▪ **SO4:** To evaluate the sensitivity and specificity of population, intervention, and outcome definitions at the patient level.
METHODOLOGY	**Study design**	Multicenter, retrospective, non-interventional study utilizing data from Electronic Health Records (EHRs) of patients in several hospitals in Spain, carried out in two sequential phases. **Phase I** will be a feasibility study and will use the project's sample size estimates to assess the adequacy of the hospital dataset for the project's aims. **Phase II** will consist of a Target Trial Emulation study comparing the per-protocol effects of full-dose long-term treatment with Bemiparin versus pooled data of Dalteparin, Enoxaparin and Tinzaparin. For each eligible patient, follow-up will begin at the time of CAT diagnosis (baseline or time zero) and will continue until the occurrence of the following: acute treatment initiation >48 hours from time zero, long-term initiation ≥20 days, treatment discontinuation or switch, VTE recurrence, death, loss to follow-up, or administrative end of follow-up (six months after baseline), whichever occurs first. Phase II of the trial would not be conducted if the objective set for Phase I is not met.
	Study periods	The study period will be from January 1st, 2014, to January 31st, 2025. The recruitment period will be from January 1st, 2015, to June 30th, 2024.
	Study population	From all patients attended at the participating hospitals (screening set), a source population will be selected using NLP filters identifying patients with cancer and a reported diagnosis of DVT, PE, or both during the recruitment period. The inclusion and exclusion criteria described below will then be applied to this source population to define the final study population. Time of CAT diagnosis in each eligible patient will be defined as time zero. Phase I will include patients from two of the participant hospitals, and Phase II will expand to the remaining.
	Inclusion and exclusion criteria	**Inclusion criteria** ▪ At least 18 years of age at diagnosis of VTE. ▪ Active cancer (Cancer diagnosis within 6 months before VTE diagnosis, treatment for cancer during the previous 6 months preceding VTE diagnosis, metastatic disease or hematological malignancy not in complete remission, all assessed at the time of CAT diagnosis). ▪ Acute episode of symptomatic, objectively confirmed DVT, PE, or both. **Exclusion criteria** ▪ Pregnancy at diagnosis of CAT. ▪ Basal cell carcinoma or non-melanoma skin cancer (in the absence of any other cancer diagnosis). ▪ VTE < 12 months before the CAT diagnosis.
	Data source	This study will be based on the secondary use of unstructured and structured data captured in the EHRs collected during the study period (January 1st, 2014, to January 31st, 2025). It will include outpatient clinic reports, discharge reports, emergency reports, prescriptions, and other medical reports.
	Data completeness assessment	To assess the completeness of the information gathered from EHRs, the total number of screened records and patients will be analyzed per site, according to the main data sources and hospital departments or services.
	Study variables	Study variables to be considered and analyzed will consist of terms recorded by physicians in the EHRs during routine clinical practice. Consequently, some variables may be excluded from the analysis if they are not present in the EHRs. Certain variables may also be inferred from directly recorded data. For a complete list of outcomes-related variables and study variables by objectives, see Sections 6.3 and 6.6, respectively. Sensitivity and specificity of population, intervention, and outcome definitions at the patient level will be performed.
	NLP data processing	The extraction of the data captured in the EHRs will be performed with EHRead®, a technology using NLP and machine learning techniques. Free-text information is translated into concepts using specific terminology (based on SNOMED-CT, ATC and LOINC) and is converted into a study

		database. External validation to evaluate the EHRead® performance will be done in each participating center.
	Data analysis methodology	Descriptive analyses and a Target Trial Emulation analysis comparing the per-protocol effects of full-dose long-term treatment with Bemiparin versus pooled data of Dalteparin, Enoxaparin and Tinzaparin. Detailed methodology for statistical analysis and a detailed description of analysis populations will be included in the statistical analysis plan (SAP).

4 BACKGROUND AND STUDY RATIONALE

Cancer-associated thromboembolism (CAT) remains a significant health issue and is one of the primary causes of mortality in cancer patients, second only to the disease itself, despite being largely preventable (1, 2). CAT typically presents as venous thromboembolism (VTE), which encompasses both deep vein thrombosis (DVT) and pulmonary embolism (PE) and may occur as either symptomatic or incidental events. This condition primarily results from the hypercoagulable state induced by cancer, driven by complex interactions between tumor cells, coagulation factors, and the immune response (3).

Individuals with cancer are at significantly higher risk of developing VTE compared to those without cancer, regardless of age (1). Over the past 20 years, the incidence of VTE in cancer patients has tripled and is now nine times greater than in the general population. (4). The reported incidence of CAT varies widely, ranging from 2% to 30%, reflecting differences in the natural history of various tumor types (5) as well as the influence of tumor-related factors (e.g., tumor location, presence of distant metastases) (5-7), treatment-related factors (e.g., surgery, chemotherapy) (8-10), and patient-related factors (e.g., older age, comorbidities, multiple surgeries) (7). Additionally, CAT patients have a two- to threefold higher mortality rate compared to those without it (4). Approximately 5.6% to 7.9% of patients with active cancer experience CAT recurrence within 6 months, being the overall prevalence of VTE in this population of 13 cases per 1000 person-year (11). From a clinical standpoint, VTE recurrence in cancer patients has been associated with increased morbidity, disease-related complications, and greater utilization of healthcare resources (12).

Current guidelines in Europe and North America recommend long-term anticoagulation therapy for patients with CAT, preferring low-molecular weight heparin (LMWH) or direct oral anticoagulants (DOACs) over vitamin K antagonists (VKA) for a minimum of six months (1, 13). The primary treatment objectives are to alleviate the acute DVT and/or PE symptoms, reduce recurrent CAT events, and decrease both fatal and non-fatal PE occurrences. However, anticoagulated cancer patients experience a higher incidence of major bleeding events, with cumulative rates of 5.9% and 8.7% at three- and six-months post-treatment initiation, respectively, compared to 2.6% and 4.2% in non-anticoagulated cancer patients (14). Consequently, before starting anticoagulant therapy in these patients, an individualized assessment of VTE recurrence and bleeding risk is crucial. This evaluation helps identify high-risk individuals and informs decisions on the choice of anticoagulant, as well as whether to extend anticoagulant therapy beyond six months for selected patients (15).

While the use of DOACs for CAT management has increased in recent years, LMWHs remain the predominant anticoagulant therapy, both in the acute phase (5–10 days) and in the long-term phase (six months). A recent study using data from the TESEO registry, promoted by the Spanish Society of Medical Oncology, found that 88% of CAT cases were treated with LMWHs, with Enoxaparin (40.8%), Bemiparin (27.0%), and Tinzaparin (20.8%) being the most commonly used (16). These findings confirm that LMWHs continue to be a cornerstone in the treatment of CAT, despite the availability of DOACs.

Notably, no data currently support the superiority of any specific LMWH over another in terms of efficacy or safety in patients with CAT (1). Clinical guidelines do not recommend one LMWH over another, and in routine practice, they are generally considered to have comparable effectiveness and safety profiles. The selection of a specific LMWH is often driven by institutional protocols, availability, and physician preference rather than by robust comparative evidence.

Bemiparin, a second-generation LMWH, has demonstrated efficacy and safety for the long-term treatment of cancer-associated VTE in routine clinical practice (17-19). However, no studies have directly compared these outcomes with those of other LMWHs, such as Dalteparin, Enoxaparin or Tinzaparin, in real-world scenarios in this population. Given the absence of evidence supporting the superiority of any specific LMWH, this study will compare long-term Bemiparin treatment with pooled data from long-term treatment with Dalteparin, Enoxaparin and Tinzaparin. This approach strengthens the study's validity by using an active comparator group with a diverse patient population, thereby improving comparability between treatment groups and enhancing the reliability of our estimates. Additionally, the diverse comparator group reduces variability, leading to a more precise assessment of Bemiparin and increasing the generalizability of the findings to different clinical settings. Medsavana's EHRead® technology enables the systematic data collection and variable extraction of a large number of EHRs across various sources and countries through natural language processing (NLP) and machine learning (ML) (12, 20-27). This technology provides a unique opportunity to analyze the effectiveness and safety of Bemiparin compared to other LMWHs, such as Dalteparin, Enoxaparin, and Tinzaparin in long-term anticoagulation in CAT management.

5 STUDY OBJECTIVES

The BEMICAT study will be carried out in two sequential phases with distinct objectives. Phase I will evaluate the feasibility of the research and if confirmed, Phase II will evaluate the effectiveness and safety of long-term Bemiparin treatment, as well as to evaluate patient characteristics and management strategies.

5.1 Phase I

To assess in a limited sample whether the incidence of CAT in patients who started long-term treatment with Bemiparin, Dalteparin, Enoxaparin, or Tinzaparin during the recruitment period is consistent with the sample size estimated for phase II.

5.2 Phase II

5.2.1 *Primary Objectives (PO)*

- **PO1**: To assess the noninferiority of long-term Bemiparin treatment in preventing VTE recurrences in adult patients with CAT over a six-month period after the index CAT episode, compared to pooled data of Dalteparin, Enoxaparin and Tinzaparin.
- **PO2:** To compare the safety of long-term Bemiparin treatment versus pooled data of Dalteparin, Enoxaparin, and Tinzaparin in adult patients with CAT over a six-month period after the index CAT episode.

5.2.2 *Secondary Objectives (SO)*

In patients with CAT on long-term anticoagulation with the drugs Bemiparin, Dalteparin, Enoxaparin, or Tinzaparin (overall and stratified by drug type):

- **SO1**: To describe the sociodemographic and clinical characteristics, including known risk factors present at the index CAT episode.
- **SO2:** To describe the patient journey and treatment pattern, including anticoagulation switches.
- **SO3**: To describe the time-to-event for effectiveness and safety outcomes.
- **SO4**: To evaluate the sensitivity and specificity of population, intervention, and outcome definitions at the patient level.

6 METHODOLOGY

To carry out the study, the Sponsor will engage the services of Medsavana and Savana Research. Medsavana, the owner of EHRead® technology, will perform all data processing. Savana Research will design the fit for purpose NLP data extraction and will perform all statistical analysis in this Study.

6.1 Study Design

This is a multicenter, retrospective non-interventional study utilizing data from EHRs of all patients with CAT attended at several hospitals across Spain from January 1st, 2015, to June 30th, 2024 (recruitment period). Both structured and unstructured real-world data (RWD) captured in the EHRs during the study period, spanning from January 1st, 2014, to January 31st, 2025, will be extracted using NLP and ML techniques (**Figure 1**). A 1-year pre-recruitment period is established, which will be necessary to apply the corresponding inclusion and exclusion criteria. Similarly, a 7-month post-recruitment period has been defined to minimize cohort effect on censoring outcomes analysis during follow-up. For all patients, time zero will be defined as the time of CAT diagnosis.

As mentioned above, the BEMICAT study will be developed in two consecutive phases (**Figure1**):

Phase I will be a feasibility study and will use the project's sample size estimates to assess the adequacy of the hospital dataset for the project's aims. It will be performed in two of the participating hospitals using data obtained from January 1st, 2014, to December 31st, 2023. The results of Phase I will be compared with the sample size estimates for the study's objectives before moving on to Phase II. For more details, please refer to sections **6.6** and **6.7** (Sample Size and Study Power and Data Analysis Plan).

Phase II will include the study population from all participating hospitals and will focus on developing the study's research objectives. The total number of hospitals to be recruited will be determined by the findings of Phase I. A Target Trial Emulation study, designed to replicate the structure and causal inference of a randomized controlled trial (RCT) within an observational dataset, will compare the per-protocol effects of full-dose long-term treatment with Bemiparin versus pooled data of Dalteparin, Enoxaparin and Tinzaparin. This approach systematically defines the eligibility criteria, time zero, interventions, and outcomes to align closely with a hypothetical RCT (28). Specifically, the study will emulate an RCT where patients with objectively confirmed CAT are assigned to long-term anticoagulation therapy

Follow-up for each eligible patient, will begin at time zero (time of CAT diagnosis) and will continue until the occurrence of any of the following events: acute treatment initiation of any anticoagulant more than 48 hours after CAT diagnosis, long-term treatment initiation ≥20 days after CAT diagnosis, treatment discontinuation or switch, VTE recurrence, death, loss to follow-up, or administrative end of follow-up (eight months after time zero), whichever occurs first. Additionally, patient-level validation will be performed to assess the sensitivity and specificity of the classification of CAT cases, the assignment to treatment subgroups of interest, and the detection of the evaluated outcomes.

Phase II of the trial would not be conducted if the objectives set for Phase I are not met.

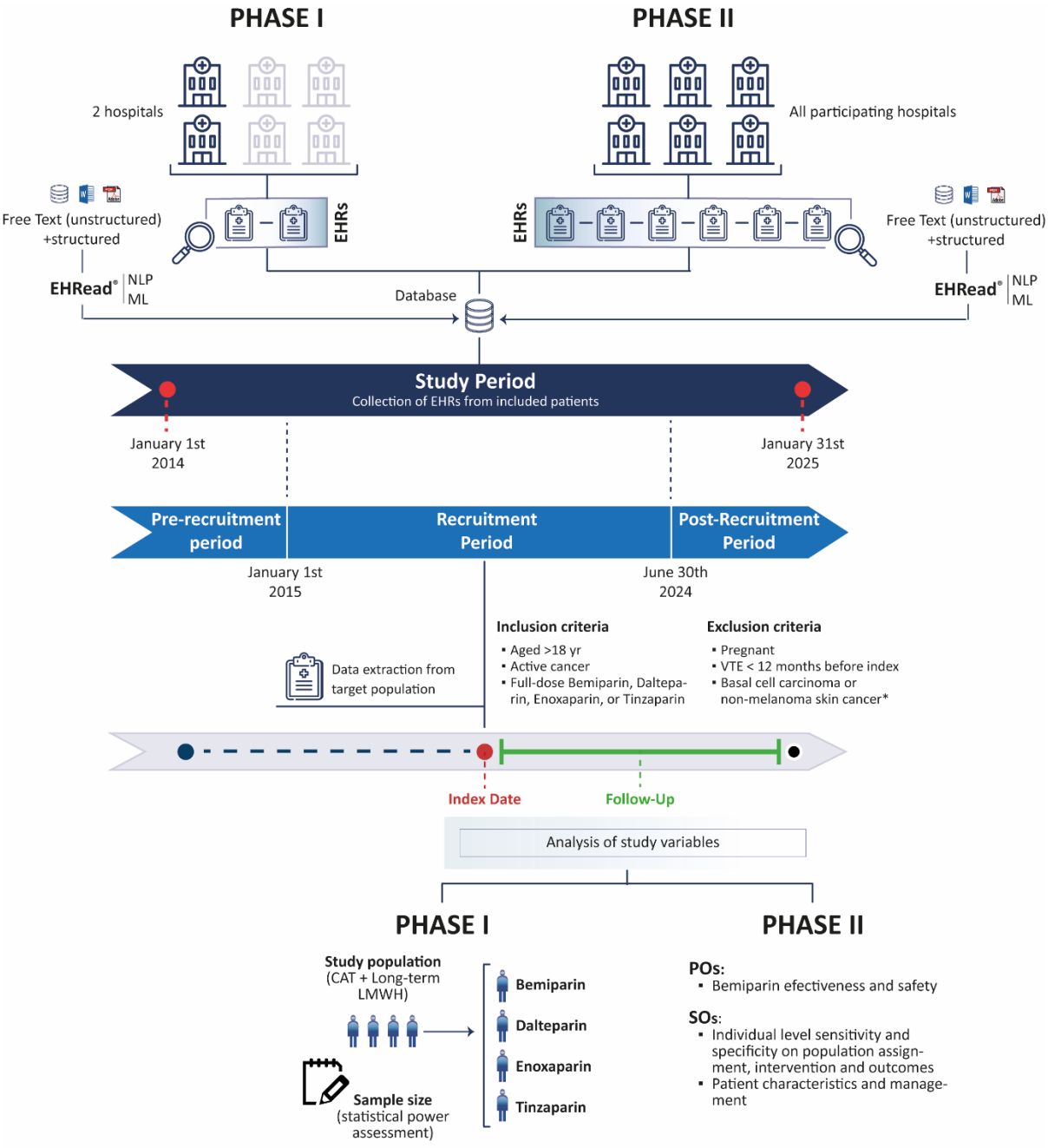

Figure 1. Study methodology. *Using the EHRead© technology, different filters for inclusion and exclusion criteria will be applied to the target population to define the study population. The clinical data from patients' EHRs will be extracted and analyzed at different time points and windows, namely time zero or Index Date and FU. The present study will comprehend two different phases. An initial Phase I will be conducted at two of the participating hospitals to assess whether the incidence of CAT in patients who started long-term treatment with Bemiparin, Dalteparin, Enoxaparin or Tinzaparin during the recruitment period is consistent with the sample size estimated for the phase II. Phase II will evaluate the effectiveness and safety of Bemiparin treatment, as well as patient characteristics and management strategies. *Active cancer will be defined as the presence of any of the following: a cancer diagnosis within 6 months before VTE diagnosis, treatment for cancer during the 6 months*

preceding the VTE diagnosis, metastatic disease, or hematological malignancy not in complete remission, all assessed at the time of CAT diagnosis. [&] DVT or PE will be considered objectively confirmed if the diagnosis is supported by imaging tests (e.g., contrast venography or compression ultrasonography for DVT; helical computed tomography scan, ventilation-perfusion lung scintigraphy, or pulmonary angiography for PE) or if the diagnostic mention in clinical notes includes terms indicating medical validation (e.g., 'confirmed DVT'). Symptomatic DVT or PE will be defined as cases where clinical signs or symptoms are documented and directly associated with the diagnosis. [#]In the absence of any other cancer diagnosis.

6.2 Study Population

For phase I, the **screening set** will include all EHRs from patients attended at two of the participating hospitals during the study period. For phase II, this will be expanded to include EHRs from patients in all participating hospitals. NLP specific filters will be applied to define the **source population**, selecting patients with cancer and a reported diagnosis of DVT, PE, or both during the recruitment period. The final **study population** will include patients with CAT from the source population adhering to the inclusion and exclusion criteria outlined below. Only patients meeting all the inclusion criteria will be eligible for participation in the study.

6.2.1 *Inclusion Criteria*

To be eligible for inclusion in the study, patients have to fulfil the following inclusion criteria:

- At least 18 years of age at diagnosis of VTE
- Active cancer*
- Acute episode of symptomatic, objectively confirmed DVT, PE, or both[&]

*Active cancer will be defined as the presence of any of the following: a cancer diagnosis within 6 months before VTE diagnosis, treatment for cancer during the 6 months preceding the VTE diagnosis, metastatic disease, or hematological malignancy not in complete remission, all assessed at the time of CAT diagnosis.

[&] DVT or PE will be considered objectively confirmed if the diagnosis is supported by imaging tests (e.g., contrast venography or compression ultrasonography for DVT; helical computed tomography scan, ventilation-perfusion lung scintigraphy, or pulmonary angiography for PE) or if the diagnostic mention in clinical notes includes terms indicating medical validation (e.g., 'confirmed DVT'). Symptomatic DVT or PE will be defined as cases where clinical signs or symptoms are documented and directly associated with the diagnosis. Specific symptoms for DVT and PE are mentioned in **Table 3**.

6.2.2 *Exclusion Criteria*

To be eligible for inclusion in the study, patients have to not meet the following criterion:

- Pregnancy at diagnosis of CAT
- Basal cell carcinoma or non-melanoma skin cancer (in the absence of any other cancer diagnosis)
- VTE within 12 months before the CAT diagnosis

6.2.3 *Definition of Subgroups in Study populations*

The study population will be categorized into the following subgroups based on the type of LMWH assigned as full-dose long-term treatment.

For **Phase II Primary Objectives**, the study population will be divided into two sub-groups:

- Bemiparin Group.
- Others LMWHs Group: Pooled Data for patients treated with Dalteparin, Enoxaparin, or Tinzaparin

For **Phase II Secondary Objectives**, the study population will be grouped as follows:

- Overall: Including the entire study population.
- Bemiparin Group.
- Dalteparin Group.
- Enoxaparin Group.
- Tinzaparin Group.
- Others LMWHs Group: Pooled Data for patients treated with Dalteparin, Enoxaparin, or Tinzaparin

6.2.4 *Criteria for Defining Full-Dose and Long-Term Treatment Strategies*

To be considered **full-dose and long-term**, treatment must meet the following conditions:

I. Fulfillment of full-dose requirements

Fulfillment of full-dose requirements will be determined by either explicit documentation in clinical notes indicating prescription at optimal doses (e.g., mentions such as "full dose", “therapeutic dose” or "complete dose") or compliance with the dosing regimens specified in section **11.1**, Annex I (17, 29).

II. Evidence of Long-Term Prescription:

The prescribed long-term treatment will be identified based on treatment entries within the reports available up to 6 weeks following CAT diagnosis. The following criteria will guide the classification of treatment strategies and censoring:

a. Treatment Strategies of Interest

- Bemiparin Group
  - Initiation of any anticoagulant within 2 days of CAT diagnosis.
  - Bemiparin started within 20 days of CAT diagnosis and continued for 6 months.
- Other LMWHs Group:
  - Initiation of any anticoagulant within 2 days of CAT diagnosis.
  - Other low-molecular-weight heparins (LMWHs) started within 20 days of CAT diagnosis and continued for 6 months.

For both groups, treatment interruption is allowed only if justified by thrombocytopenia or an invasive procedure, with an interruption not exceeding 2 weeks.

b. Assignment and Censoring Protocol

At the time of CAT diagnosis, patients will be provisionally assigned to both treatment strategies. A treatment strategy will be considered incompatible if the actual anticoagulation course deviates from the predefined criteria for initiation or maintenance, leading to patient censoring due to non-adherence to the assigned strategy.

c. Stepwise Censoring Rules

- Step 1: Initial Anticoagulant Use
  - If no anticoagulant is initiated within 2 days of CAT diagnosis, the patient will be censored from both strategies.
- Step 2: Drug Documentation (3–20 Days)
  - Patients without documented drug entries between 3- and 20-days after CAT diagnosis will be censored from both strategies.
  - If a single drug is documented within 3–20 days after CAT diagnosis, and at least one entry occurs between 10 and 20 days, the patient will remain assigned to the compatible treatment strategy and will be censored from the incompatible treatment strategy.
  - If multiple drugs are documented within 3–20 days after CAT diagnosis, and at least one entry occurs between 10 and 20 days, the patient will remain assigned to the treatment strategy aligned with the last drug documented in this timeframe and will be censored from the incompatible treatment strategy.
- Step 3: Treatment Maintenance
  - Any cessation of anticoagulation will result in censoring, except when justified by thrombocytopenia or an invasive procedure, provided that the interruption does not exceed 2 weeks.

Table 2 outlines the outcomes of the study.

Table 2. Outcomes.

Outcomes	Definition	Timing	Notes
Effectiveness outcomes			
Primary Outcome:			
▪ Symptomatic VTE recurrence[&]	VTE recurrences is defined as a new episode of symptomatic and objectively confirmed DVT, fatal or nonfatal PE, or synchronic DVT and PE occurring during long-term treatment or within one month after the end of treatment.	Follow-up*	Recurrent DVT will be considered as evidence of thrombus in the same vein without signs of chronic thrombosis changes, or as a new episode in a different location. Recurrent PE will be considered as new embolus in the pulmonary arteries, excluding residual or chronic thromboembolic changes, or as a new PE occurring in patients receiving optimal-dose anticoagulation therapy for a prior DVT. Fatal PE will be identified either through explicit physician reports attributing death to PE or by a temporal confluence of PE and death under pre-established criteria (e.g., a confirmed PE diagnosis within 48–72 hours prior to death and no other evident cause) Synchronic DVT and PE refer to simultaneous episodes of DVT and PE or DVT followed by PE within 7 days; if both occur but do not meet these criteria, only the episode closest to the initiation of long-term treatment will be counted as the outcome.
Secondary Outcomes:			
▪ Incidental VTE recurrence	New episodes of objectively confirmed[#] DVT or fatal and nonfatal PE occurring during long-term treatment or within one month after the end of treatment, in the absence of clinical signs or symptoms suggestive of DVT or PE.	Follow-up*	The criteria for recurrent DVT, PE, or synchronic DVT and PE, defined for symptomatic VTE, also apply to incidental cases.
▪ Stroke	Episodes of stroke during long-term treatment or within one month after the end of treatment	Follow-up*	NA
▪ Acute myocardial infarction	Episodes of acute myocardial infarction diagnosed during long-term treatment or within	Follow-up*	NA

	one month after the end of treatment.		
▪ All- cause Death[&] (30)	Documented evidence of mortality from any cause, either directly reported or indirectly inferred from terms that clearly indicate the patient's death (e.g., "end-of-life situation," "comfort measures initiated")	Follow-up*	This outcome will be limited to deaths recorded in the hospital due to the lack of access to external mortality databases. Mortality inferences will be detailed in the SAP.
Safety outcomes			
Primary Outcome:			
▪ Major bleeding (fatal or non fatal)[&] (30, 31)	Major bleeding is defined as any bleeding event that meets at least one of the following criteria during Long-term treatment or within 24 hours after the end of treatment: ▪ Fatal bleeding ▪ bleeding occurring in a critical location (e.g., intracranial, intraocular, intraspinal, pericardial, retroperitoneal, intraarticular, or intramuscular with compartment syndrome). ▪ Clinically overt bleeding associated with a decrease in hemoglobin of ≥2 g/dL (20 g/L; 1.24 mmol/L) compared to the pre-bleeding level. ▪ Clinical overt bleeding requiring a transfusion of ≥2 units of red cells or whole blood ▪ Clinically overt bleeding necessitating surgical intervention.	Follow-up*	Only the major bleeding event detected closest to the initiation of long-term treatment will be considered as the primary outcome. Any subsequent major bleeding events, regardless of their severity or timing, will not be included in the outcome analysis.

Secondary Outcomes			
▪ Clinically relevant non-major bleeding (CRNMB) & (30, 32)	Any bleeding recorder during treatment or within 24 hours after the end of treatment that does not meet the criteria for major bleeding but meeting at least one of the following criteria: ▪ Requires medical attention, such as hospitalization, medical treatment for bleeding, or an increased level of care. Prompts a face-to-face evaluation in an outpatient setting ▪ Leads to a change in antithrombotic therapy (including discontinuation or down titration of study drug)	Follow-up*	As with major bleeding, only the earliest CRNMB event following the initiation of long-term treatment will be considered in the analysis. However, major bleeding and CRNMB events will not be mutually exclusive, allowing for both types of events to be independently recorded and analyzed, even if they occur in the same patient.
▪ Clinically relevant bleeding (30)	Defined as the rate of patients experiencing at least one major bleeding and/or a clinically relevant non-major bleeding	Follow-up*	NA
Treatment-related outcomes			
▪ Treatment patterns	▪ Total count and frequency of distinct anticoagulant received at diagnosis of CAT (acute phase) ▪ Total count and frequency of distinct anticoagulant received at time zero (long-term) ▪ Total count and frequency of distinct anticoagulant received during the follow-up ▪ Number of patients with long-term treatment discontinuation.	Pre-time zero and follow-up*	Discontinuation is defined as any cessation of treatment, except when caused by thrombocytopenia or an invasive procedure, followed by reinitiation within less than 2 weeks.
▪ Treatment switch	Reported evidence of any anticoagulant other than the treatment assigned as long-term.	Follow-up*	To confirm the switch, the new treatment must be documented in at least two consecutive reports. Only the first switch recorder will be counted as outcome. The date when the new anticoagulant is initiated will be switch date.

*[&] At least two records from any of the following: Oncology, Internal Medicine, Hematology, Thrombosis units or Emergency Departments, will be required for confirmation. * Follow-up: From time zero until VTE recurrence, treatment discontinuation or switch, death, loss to follow-up, or the administrative end (six months), whichever occurs first. [#] The same criteria for defining the index VTE case will apply: DVT or PE recurrence is confirmed if supported by imaging or explicitly validated in clinical notes, while symptomatic DVT or PE requires documented signs or symptoms linked to the diagnosis.*

6.3 Data Source

The data source will be the free-text, unstructured and structured information in patients' EHRs, from several hospitals within the Spanish Healthcare Network (**Figure 1**). It will include outpatient clinic reports, discharge reports, emergency reports, prescriptions, and other medical reports. Images, such as hand-drawn pictures and scanned images, will not be extracted. No data entry by physicians or their delegates into an electronic data capture platform will be performed. No Clinical Research Documents will be collected for this study. Electronic data will be collected from all available services and departments in each participating site, including emergency, external consultations, and hospitalization notes.

6.4 Data Processing

6.4.1 *Data Acquisition*

The Data Acquisition Phase is the responsibility of the participating site (IT service) alongside a close collaboration with Medsavana Information Technology (IT) Staff. However, specific documentation listing the necessary data sources to construct the study database will be created based on this protocol. Briefly, the list of data sources contains information about the necessary type of data (structured, semi-structured, free text, etc.), hospital departments, and hospital areas (emergency room, outpatients) needed to assess the objectives of the study. Additionally, if required by one or more participating hospitals, the list of data sources will include a comprehensive list of ICD-9/ICD-10 codes corresponding to the patients with the diseases of interest to the study. Data acquisition will be initiated only after the conclusion of the study period (February 2025) and will be conducted exclusively in a retrospective manner, with no additional data extractions planned.

6.4.2 *Data Integration*

Medsavana, in close collaboration with participating site IT service, will receive the EHRs from heterogeneous sources, as every site may have different Information Systems. This information is then uploaded to a secure file transfer protocol utility exclusively available for each site. In the integration stage, the EHRs will be included in an inventory to be prepared for the NLP phase (*EHRead*® technology). This step comprises format standardization, data cleaning, data quality reporting, and application of business rules.

6.4.3 *Data Completeness Assessment*

Once participating hospitals have been signed and their data is integrated, a first data quality report will be generated to assess the completeness of the information gathered from EHRs focusing especially on study key departments and report types. The total number of screened records and patients will be analyzed per site, according to the main data sources (admission, consultation, or emergency notes) and hospital department or services (Oncology, Hematology, Thrombosis units, Vascular Surgery, Internal Medicine, Pneumology, Emergency department, and others). Only those hospitals that fulfill all the criteria described will finally enter the study.

6.4.4 *NLP Data Processing to Generate Study Database*

6.4.4.1 Data extraction and EHRead® Technology Overview

Medsavana developed the EHRead® technology, an integrated suite of diverse Python-based modules that leverage various computational techniques, with a strong emphasis on NLP, for extracting clinically relevant information from EHRs' free text and providing the information in a structured database (22, 27, 33-38). EHRead® is capable of meaningfully interpreting the content included in the free text in various languages (Spanish, Catalan, English, French or German), regardless of the EHR system on which it operates. It processes both structured data and free text information from EHRs detecting clinical entities and mapping them to standard terminologies, such as Systematized Nomenclature of Medicine (SNOMED CT), Anatomical Therapeutic Chemical (ATC) and Logical Observation Identifiers Names and Codes (LOINC) classifications (39, 40). SNOMED CT provides codes, concepts, synonyms, and definitions for clinical documentation, covering symptoms, diagnoses, body structures, and substances, and is considered the most comprehensive terminology in the world (40). The ATC system categorizes drugs based on their therapeutic, pharmacological, and chemical properties, while LOINC uniquely codes laboratory tests across various domains, offering a standardized approach for reporting and exchanging results across different healthcare systems. This approach enables the reuse of information included in large-scale collections of clinical records. *EHRead®'s* modules are integrated into a high-throughput big data processing pipeline, utilizing highly parallelized cloud computing with Spark. While the overall pipeline is proprietary, ML models may incorporate open-source architectures such as transformers, and models are fine-tuned for each use case to ensure optimal performance (see **6.4.4.2**).

In this study, *EHRead®* will primarily be applied to the dataset obtained from the participating hospitals (see data acquisition section), and the final study population will be determined through the application of NLP filters. To comply with the principle of data minimization, only patients who fulfil all inclusion and no exclusion criteria will be included in the final study database which serves as the foundation for subsequent statistical analyses.

Data extraction by the use of *EHRead®* will be carried out by the participating site and Medsavana IT staff.

6.4.4.2 Model Development

The NLP models used in *EHRead®* are iteratively optimized in a robust machine learning lifecycle to ensure optimal performance. For each specific study, during the NLP development phase, study-specific terms are integrated, and model performance is evaluated. If model drift is detected, retraining and model update is necessary, proceeding later with the study data processing.

General and specific models will be used in this study and key processing steps include the following:

General NLP models

- Study data capture: named-entity recognition and linking (NER-NEL) models identify clinical entities and link them with standard terminology (e.g., SNOMED CT, ATC and LOINC)
- Negation detection: clinical entities are classified as 'affirmative' and 'non-affirmative'.
- Copy paste deletion: duplicate text within patient documents is flagged and taken as input for downstream analysis.
- Section detection: paragraphs, sentences and sub-sentences of the EHR are categorized based on clinical context (family background, patient background, or patient present).
- Temporality: temporal relationships are established, linking clinical entities to time references such as absolute/explicit dates and times, as well as time differences ("last week", "two months ago", etc.).
- Measurable parameters: Quantitative clinical data, such as test results (e.g., laboratory magnitudes) or biometric variables (e.g., body-mass index), are extracted and standardized for consistent analysis through different hospitals and services.

Specific NLP models

They are tailored to the study needs and include the extraction of any kind of clinical entity and/or attribute considered relevant to answer the study objectives. For its development, trained medical research experts conduct dedicated annotation projects to generate the necessary training corpora. The performance of those models is also evaluated.

6.4.5 *External Validation of EHRead® performance*

6.4.5.1 Internal Validation

To ensure high-quality NLP processing, Medsavana's medical research experts will review term detection summaries (result from the NER-NEL model application) for false negatives and false positive detections across all study languages.

To identify potential false negatives of variables considered of key/critical importance for a study, a sample of EHRs enriched with mentions of critical variables, focusing on study population definitions, analysis subgroups, treatments, and outcomes, will be reviewed by expert annotators. A published methodology (22) determines the minimum EHR sample size required for review, selecting representative documents based on variable frequency. During the review, potential false negatives are flagged and subsequently validated by a team of medical research experts. Once confirmed, these cases are incorporated into the retraining of the NER-NEL module to improve its performance.

After initial training has been completed and NER-NEL has been applied to the study database, representative examples for each relevant clinical term, including all alternative names (i.e., synonyms) and acronyms, are retrieved and reviewed by Medsavana's annotators to identify potential false positives, which will be subsequently evaluated by medical experts, who must modify terminology accordingly for subsequent retraining of the NER-NEL module.

Once the initial training phase has concluded, the effectiveness of this process will be finally assessed by standard performance metrics for key/critical study variables: precision, recall, and their harmonic mean, or "F1-score" (41). The metrics are defined as follows:

Precision = $\frac{tp}{tp+fp}$ Indicates the accuracy of the information the system retrieves.

Recall = $\frac{tp}{tp+fn}$ Indicates the amount of information the system retrieves.

F1-Score = $\frac{2\times Precision\times Recall}{Precision+Recall}$ Overall performance indicator of information retrieval.

In all cases, *tp* represents the number of true positives, *fn* represents the number of false negatives (i.e., records incorrectly not retrieved), and *fp* represents the number of false positives (i.e., records incorrectly retrieved).

If the F1-score for any key or critical variable is below 0.9, the process of annotation, terminology update, retraining, and validation will be repeated until this minimum threshold is achieved. If this threshold cannot be achieved due to limitations inherent in the data or technical constraints, this will be explicitly reported, along with a discussion of the potential implications for the analysis.

6.4.5.2 External Validation

The clinical findings will be complemented with an evaluation of the performance of *EHRead®* by external investigators participating in the research study. This evaluation aims to verify the NLP performance when identifying records that contain critical variables for the pathology of interest and its related clinical context (42) ensuring the consistency of the guidelines and the reliability of the annotated parameters (43). It is important to note that it will be assumed that the subset of patients identified with registered or available data regarding the pathology of interest will be representative of the disease, and no corrections will be applied to the population sample. This evaluation will be performed at each participating site of the study. A random sample of the EHRs annotated internally for performance evaluation (see section **6.4.5.1**) will be selected and made available to local annotators at each participating site.

The workflow will be as follows:

- Two designated physicians (hereinafter referred to as "the annotators") at each hospital participating in the study will annotate a set of randomly selected records from their site. These physicians will be selected by the Principal Investigator at each site. They will annotate clinical entities relevant to the study. The annotators will not be allowed to communicate with each other, as far as the annotation process is concerned, but they could turn to the medical team of Medsavana when needed.

- Once the annotations are finalized, a third physician from the same hospital will assess them, reviewing the annotations made by the two annotators and resolving differences, yielding the gold standard corpus to evaluate *EHRead®* performance (**Figure 2**).

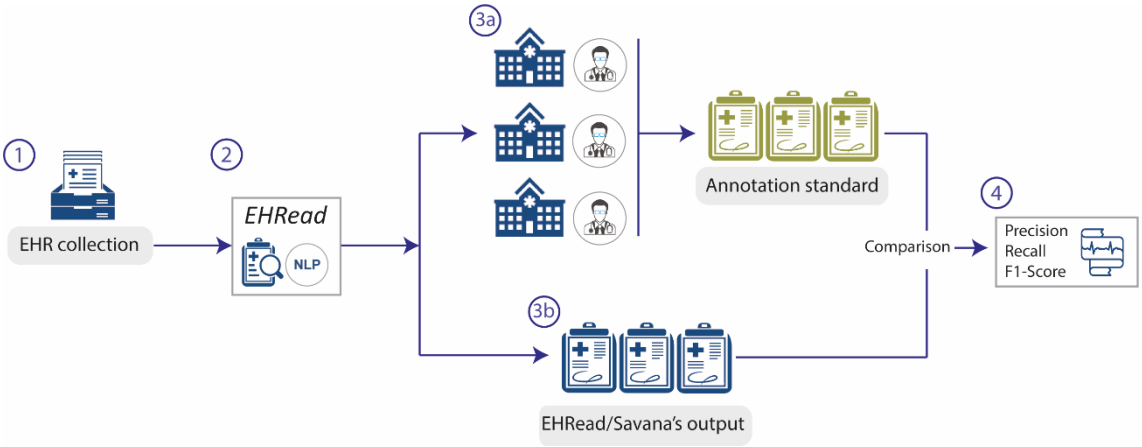

Figure 2. Evaluation Process. *A set of EHRs (1) collected via EHRead© (2) is annotated by physicians to generate the annotation standard (3a) and compare it to the study results (3b), obtaining precision, recall, and the F1-scores, to evaluate the quality of the Medsavana's output (4)*

Results of this annotation process are the inter-annotator agreement (IAA) and the performance of the *EHRead©* system as reported using the standard metrics of precision, recall, and F1-score", as previously described.

These metrics will be reported in the final analysis as an average of the metrics across all hospitals. For any variable, were the external F1-score deviated by more than >15% from the internal F1-score, a dedicated investigation will be conducted to identify the cause of the deviation. The gold standard for resolving disagreements will be the annotation results from the internal validation project. This process will be documented in the Clinical Scientific Report (CSR).

6.5 Study Variables

The list of variables of interest will be specifically generated or constructed for this study and will be clinically supervised by Medical Research Experts. The study variables considered and analyzed will be collected from the unstructured and structured information in the EHRs registered by the physicians during routine clinical practice. The study variables related to demographic characteristics, such as age and sex, the patient journey-related variables, such as department visits and hospitalizations, laboratory data, and pharmacy data will be extracted from structured text. Variables related to comorbidities, clinicopathological characteristics, disease outcomes, and image test and procedures will be extracted from unstructured text.

The list of variables of interest in the present study are shown in **Table 3** by objective (Primary objective: PO; Secondary objectives: SO-1, SO-2...), classified by patient population, category, and time point for analysis (TZ: time zero; FU: Follow Up). As previously stated, the study variables and the terms that compose them (which will be included in the SAP), are collected from the information registered by the physicians, therefore, it is possible that some of the variables below will not be included in the analysis if they are not contained in the EHRs. The study variables will be analyzed at time zero and follow-up, whenever applicable, considering a time window around each timepoint to optimize data collection (to be specified in the Statistical Analysis Plan; SAP; See Section **6.7 Data Analysis Plan**).

Table 3. List of study variables.

Primary Objective (PO)			
Patient population	**Variable category**	**Variable/Endpoint**	**Time point**
PO1: To assess the noninferiority of long-term Bemiparin treatment in preventing VTE recurrences in adult patients with CAT over a six-month period after the index CAT episode, compared to pooled data of Dalteparin, Enoxaparin and Tinzaparin.			
Study population LMWHs subgroups	*Primary effectiveness outcomes*	▪ Symptomatic VTE recurrence ○ Symptomatic recurrent DVT ○ Symptomatic recurrent non-fatal PE ○ Fatal PE ○ Synchronic symptomatic DVT and PE	FU
	Secondary effectiveness outcomes	▪ Incidental VTE recurrence ○ Incidental recurrent DVT ○ Incidental recurrent non-fatal PE ○ Incidental Synchronic DVT and PE ▪ Stroke ▪ Acute myocardial infarction ▪ All-cause mortality.	
PO2: To compare the safety of long-term Bemiparin treatment versus pooled data of Dalteparin, Enoxaparin, and Tinzaparin in adult patients with CAT over a six-month period after the index CAT episode.			
Study population LMWHs subgroups	*Primary safety outcome*	▪ Major bleeding	FU

Secondary outcomes	*safety*	▪ CRNMB ▪ Clinical relevant bleeding

For PO1 and PO2 potential confounders from variables defined in SO1, SO2 and SO3 will be evaluated at TZ and during the FU.

Secondary Objectives (SO)			
Patient population	**Variable category**	**Variable/Endpoint**	**Time point**
SO-1 In patients with CAT on long-term anticoagulation with the drugs Bemiparin, Dalteparin, Enoxaparin or Tinzaparin (overall and stratified by drug type), to describe the sociodemographic and clinical characteristics, including know risk factors present at the index CAT episode.			
Study population LMWHs subgroups	*Sociodemographic characteristics*	▪ Age ▪ Race ▪ Sex ▪ Smoking ▪ Alcohol abuse	TZ*
	Anthropometric characteristics	▪ Weight ▪ Height ▪ BMI	TZ
	Comorbidities and antecedents	▪ Cardiovascular disease o Hypertension o Dyslipidemia o Ischemic heart disease o Heart failure o Atrial fibrillation o Peripheral arterial disease o Varicose veins o Stroke o PE o VTD ▪ Autoimmune diseases o Inflammatory bowel disease o Vasculitis (e.g., Behçet) o Antiphospholipid syndrome o Others (e.g., rheumatoid arthritis, systemic lupus erythematosus) ▪ Metabolic diseases o Obesity o Metabolic syndrome o Diabetes mellitus ▪ Myeloproliferative syndromes ▪ Monoclonal gammopathy ▪ Nocturnal paroxysmal hemoglobinuria ▪ Anemia or transfusions ▪ Previous bleeding	TZ

	▪ Peptic ulcer ▪ Intestinal angiodysplasia ▪ Colonic polyposis ▪ Psychiatric disorders ▪ Chronic kidney disease ▪ Chronic obstructive pulmonary disease ▪ Asthma ▪ Chronic liver disease ▪ Portal hypertension ▪ Genetic factors o Congenital thrombophilia o Deficiency of: - Protein C - Protein S - Antithrombin - Factor V Leiden - Prothrombin G20210A	
Concomitant treatments	▪ Antihypertensive drugs ▪ Lipid-lowering agents ▪ Antidiabetic treatments ▪ Antidepressant ▪ Antiplatelets agents ▪ Erythropoietin ▪ Hormonal contraceptives ▪ Corticotherapy ▪ Hormone replacement therapy ▪ Heparin ▪ Antipsychotics	TZ
Mechanical factors	▪ Venous catheter ▪ Other endovascular devices ▪ Immobility ▪ Recent surgery	TZ
Cancer characteristics	▪ Location ▪ Time since cancer diagnosis ▪ TNM stage ▪ Histology ▪ Molecular alterations ▪ ECOG PS ▪ Cancer status (e.g, partial response, progression)	TZ
Cancer therapy	▪ Radiotherapy ▪ Chemotherapy ▪ Hormonal ▪ Immunotherapy ▪ Targeted therapies	TZ
Laboratory values	▪ Hemoglobin ▪ Platelet count ▪ Leucocytes	TZ

		▪ Coagulation times	
	VTE-related variables	▪ Time from cancer diagnosis to VTE diagnosis ▪ Symptoms of DVT (e.g.,limb pain, skin discoloration) ▪ Symptoms of PE (e.g., dyspnea, chest pain)	TZ
SO-2: In patients with CAT on long-term anticoagulation with the drugs Bemiparin, Dalteparin, Enoxaparin, or Tinzaparin (overall and stratified by drug type), to describe the patient journey and treatment pattern, including anticoagulation switches.			
Study population LMWHs subgroups	*Patient journey*	▪ Department of CAT diagnosis. ▪ Departments visited ▪ Hospitalizations ▪ Emergency room visits	TZ, FU
Study population	*Treatment patterns*	▪ Total count and frequency of distinct anticoagulant received at diagnosis of CAT (acute phase, long-term, overall): ○ Bemiparin ○ Dalteparin ○ Enoxaparin ○ Tinzaparin ○ Other LMWHs ○ Non-LMWHs anticoagulants	
LMWHs subgroups	*Treatment patterns*	▪ Number of patients with long-term LMWHs discontinuation	
SO-3: In patients with CAT on long-term anticoagulation with the drugs Bemiparin, Dalteparin, Enoxaparin, or Tinzaparin (overall and stratified by drug type), to describe the time-to-event for effectiveness and safety outcomes.			
Study population LMWHs subgroups	*Effectiveness outcomes*	▪ Time to VTE recurrence ○ Time to recurrent DVT ○ Time to recurrent non-fatal PE ○ Time to recurrent fatal PE ○ Time to synchronic DVT and PE ▪ Time to stroke ▪ Time to acute myocardial infarction	FU
	Safety outcomes	▪ Time to first major bleeding ▪ Time to first CRNMB ▪ Time to clinically relevant bleeding	FU
SO-4: In patients with CAT on long-term anticoagulation with the drugs Bemiparin, Dalteparin, Enoxaparin, or Tinzaparin (overall and stratified by drug type), to evaluate the sensitivity and specificity of population, intervention, and outcome definitions at the patient level.			
Study population LMWHs subgroups	*Sensitivity and specificity*	▪ CAT cases ▪ Long-term LMWH assigned ▪ VTE recurrence ▪ Major bleeding	FU

		▪ CRNMB ▪ All-cause death	

**Variables will be searched in all EHRs available within a specified time-frame window relative to pre- and post-time zero if not present at this timepoint. Variations of the length of this time-frame window for some variables will be specified in the Statistical Analysis Plan. TZ: Time zero; FU: Follow Up.*

6.6 Sample Size and Study Power

The estimated number of patients with CAT receiving long-term Bemiparin treatment recruited is approximately 1,640, while other LMWHs treatment recruited is approximately 4,268. This estimate is based on: i) a theoretical population at risk of 2,400,000; iii) a recruitment period of 9.5 years (from January 1, 2015, to June 30, 2024); iv) an incidence rate of 9.5 per 100,000 person-years for patients with CAT receiving long-term Bemiparin treatment (16); and v) an anticipated 20% loss due to missing data or incomplete treatment information.

The estimated statistical power of this project to assess the non-inferiority risk ratio for VTE recurrence between Bemiparin and other LMWHs long-term treatments over 6 months exceeds 0.91 -99. This estimate is based on: i) a total of 1,640 patients with CAT receiving Bemiparin long-term treatment; ii) a total of 4,268 patients with CAT receiving other LMWHs long-term treatment; iii) a 6-month cumulative VTE recurrence rate of 2.1% for patients with LMWHs long-term treatment (29); and iv) a non-inferiority margin of 1.5 (29). Additionally, the minimum number of patients with CAT receiving long-term treatment with Bemiparin required to achieve a statistical power of 0.80, assuming all other factors remain constant, is 1,052.

The estimated statistical power of this project to assess the 6-month risk ratio of major bleeding between Bemiparin and other low molecular weight heparins (LMWHs) exceeds 0.96. This calculation is based on the following assumptions: i) a total of 1,640 patients with CAT receiving long-term Bemiparin treatment; ii) a total of 4,268 patients with CAT receiving other LMWHs long-term treatment; iii) 6-month cumulative major bleeding rates of 1.3% for patients treated with Bemiparin and 2.9% for those treated with other LMWHs (17,39); and iv) an alpha risk of 0.05. Additionally, the minimum number of patients with CAT receiving long-term treatment with Bemiparin required to achieve a statistical power of 0.80, assuming all other factors remain constant, is 798.

6.7 Data Analysis Plan

6.7.1 *Statistical Analysis Framework*

A comprehensive description of the study population and methodology for statistical analyses will be outlined in the SAP. The SAP may lead to minor modifications of the plans outlined in the protocol; any major modifications of primary endpoint definitions or their analyses will be reflected in a protocol amendment.

Specific analysis to accomplish the objectives of Phase II are detailed below:

6.7.2 *Analysis Plan per Objectives*

Phase I *(To assess in a limited sample whether the incidence of CAT in patients who started long-term treatment with Bemiparin, Dalteparin, Enoxaparin, or Tinzaparin during the recruitment period is consistent with the sample size estimated for the phase II)*. To conduct this assessment, the annual incidence of patients with CAT receiving long-term treatment with Bemiparin will be estimated in the two hospitals included in this phase. A predictive model, adjusted for the interaction between hospital type, hospital population, and year, will be used to estimate the annual incidence per hospital-year across all the hospitals participating in the study. The adequacy of this estimate will be evaluated by ensuring that its confidence intervals encompass the expected 7.5 cases per 100,000 persons-year (see section **6.6** Sample Size and Study Power).

Additionally, based on the participant loss observed by Trujillo-Santos et al., it was determined that for progression to phase II, with sufficient confidence to achieve a study power of at least 0.80, the confidence interval of the estimate must not include the minimum required sample size plus a 45% increase. This corresponds to 5.7 cases per 100,000 persons-year (see section **6.6**, second paragraph, 'Sample Size and Study Power').

Phase II Primary Objectives:

PO-1: *To assess the noninferiority of long-term Bemiparin treatment in preventing VTE recurrence in patients with CAT over a six-month period after the index CAT episode, compared to pooled data of Dalteparin, Enoxaparin, and Tinzaparin* ***and PO-2:*** *To compare the safety of long-term Bemiparin treatment versus pooled data of Dalteparin, Enoxaparin, and Tinzaparin in adult patients with CAT over a six-month period* after *the index CAT episode.*

To achieve these objectives, the analysis will focus on the per-protocol effect, following a target trial emulation framework, which applies randomized trial design principles to observational studies. By explicitly defining the hypothetical trial, this approach enables causal inference and mitigates biases such as confounding and improper follow-up specifications, ensuring the study design closely mimics a RCT (28). Parametric estimators will be used for the time-to-event outcomes, specifically using weighted pooled logistic regression to compare risks between groups. A noninferiority margin of 1.5, based on the findings of Trujillo-Santos et al (29), is selected. Noninferiority of Bemiparin compared to a pooled data of Dalteparin, Enoxaparin, and Tinzaparin in terms of VTE incidence will be established if the upper bound of the 95% confidence interval (CI) for the risk ratio is below 1.5. The Wald test will be used to assess the statistical significance of this result.

Before estimating the weighted pooled risks between groups using logistic regression, several steps will be undertaken. First, a cloning procedure will be applied to compare sustained treatment strategies which are not uniquely defined at zero time. This ensures that individuals are assigned to treatment strategies based solely on pre-time zero information, thereby preventing immortal time bias. Second, individuals who deviate from their assigned treatment strategy will be censored, ensuring that only the clones adhering to their assigned strategy remain under follow-up. However, this approach may introduce selection bias during follow-up. Third, to address this selection bias, time-varying inverse probability weights will be estimated from time zero to censoring during follow-up.

These weights redistribute the influence of censored individuals to those who remain uncensored, ensuring an unbiased representation of treatment effects over time. Potential effect modifiers (e.g, index DVT or PE) will be part of model specification to adjust the comparison of average treatment effects across treatment strategies.

Phase II Secondary Objectives:

S0-1 *To describe the sociodemographic and clinical characteristics at index CAT episode,* and **SO-2** *to describe the patient journey and treatment pattern, including anticoagulation switches*

Descriptive statistics will be used to report patient demographics, clinical characteristics, switch rates and treatment patterns. Frequency tables will be generated for categorical variables, whereas continuous variables will be described by means of summary tables that may include the mean, standard deviation, median, minimum, maximum, and quartiles of each variable. The number of non-evaluable outcomes and of missing data will also be provided and will not be counted in the percentages.

SO-3: *To describe the time-to-event for effectiveness and safety outcomes*.

Time-to-event for the evaluated outcomes of interest will be assessed using a Cumulative Incidence Function (CIF). CIF is chosen over Kaplan-Meier as it provides a clear visual representation of how the probability of the event evolves over time, considering competing risks like death. The cumulative incidence plot will represent the probability of event occurrence over the 6-month period for each group (e.g., Bemiparin vs pooled data of Dalteparin, Enoxaparin and Tinzaparin), with confidence intervals and the number at risk at relevant time points.

S0-4: *To evaluate the sensitivity and specificity of population, intervention, and outcome definitions at the patient level.*

The evaluation process for validating definitions generated by NLP algorithms and clinical rules for the population, intervention, and outcome involves three main steps: generating an independent sample, establishing a reference standard, and comparing the reference standard to the generated definitions for the independent sample.

For population validation, an independent sample will be selected using primarily structured data to increase the frequency of the population of interest. For example, to generate a sample for CAT cases, patients will be selected if they have at least two visits to the medical oncology department and at least one prescription of anticoagulant treatment. This approach aims to achieve approximately 50% representation of CAT cases. The sample size for categorical variables will be calculated following the methodology proposed by Hajian-Tilaki et al. (44), while for continuous variables, it will be determined based on the Bland-Altman limits of agreement, following the approach proposed by Lu et al. (45).

The reference standard will be established by an independent and blinded Clinical Adjudication Committee (CAC) composed of three medical experts from Medsavana. These physicians will not participate in the analysis or interpretation of the study results. Furthermore, they will remain blinded to the rules governing the NLP algorithms and clinical rules. Patients' adjudication will be conducted

independently by two medical experts, with the third serving as a reviewer in the event of disagreement. If consensus cannot be reached, a formal consensus meeting will be convened. The adjudication process will adhere to an adjudication charter developed in accordance with the methodology outlined by Kradjian et al. (46). Additionally, interim analyses will assess both intra- and inter-reviewer agreement.

To compare the reference standard with the generated definitions, classification statistics will be employed for categorical variables, including the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). These metrics will assess the validity of the generated definitions in correctly detecting true positives and true negatives relative to the reference standard. For continuous variables, agreement will be assessed using intraclass correlation coefficient (ICC) estimates.

For outcome and intervention validation, the population of interest from initial sample (~50% of initial sample) will be expanded using the validated population definitions. This will be done to reach the sample size required by the calculations for outcome and intervention validation. The expanded population will undergo a verification process to eliminate all false-positive cases, ensuring that only true cases are included. The sample size, the reference standard, and comparisons between generated definitions and reference standard will be executed using the same methodology as for population validation(45).

Lastly, for subsequent study aims, if the desired specificity (≥90%) for categorical variables or agreement (ICC ≥0.90) for continuous variables is not achieved, an additional verification process will be implemented. This process will ensure the elimination of all false-positive cases from the study dataset.

6.7.3 *Missing data handling*

Missing data will be handled according to the nature of the data collection process (see description of the *EHRead®* technology in Section **6.4.4**), namely assuming that physicians reflect clinically relevant information in EHRs. In this context:

- For Boolean variables such as comorbidities or symptoms, missing data will be assumed as true absence (i.e., the patient lacks that comorbidity/symptom). For other dichotomic variables that do not fulfil this criterion (e.g., treatment response) their absence will not be imputed and the number of patients with missing data for those variables will be reported.
- For multi-level (≥3) categorical variables and specific variables such as lifestyle factors (e.g., smoking or drinking habits), their absence will not be imputed. Instead, the number of patients with missing data will be classified into a separate category and remain as part of percentage calculations.
- For numeric variables, missing values will not be imputed. Summary statistics will be calculated based only on available (non-missing) values, and the proportion of patients with missing data will be shown.

6.8 Quality Control and Quality Assurance

This study will be executed under the framework of this study protocol and a SAP to guarantee that the data are "fit-for-purpose" and that the analysis approach is pre-specified and systematic. In this context, several quality checks (**Figure 3**) will be conducted throughout the project to ensure the following: 1) designing the data extraction strategy from each hospital (data completeness and consistency described in sections **6.4.1**, **6.4.2**, and **6.4.3**), 2) engineering the NLP processing (internal and external *EHRead©* performance evaluation described in section **6.4.5.1** and **6.4.5.2**) and 3) executing the appropriate analyses to meet the study objectives (analysis evaluation described in section 6.8 and fully described in SAP).In this context, several quality checks (**Figure 3**) will be conducted throughout the project to ensure the following: 1) designing the data extraction strategy from each hospital (data completeness and consistency described in sections **6.4.1**, **6.4.2**, and **6.4.3**), 2) engineering the NLP processing (internal and external *EHRead©* performance evaluation described in section **6.4.5.1** and **6.4.5.2**) and 3) executing the appropriate analyses to meet the study objectives (analysis evaluation described in section **6.7** and fully described in SAP).

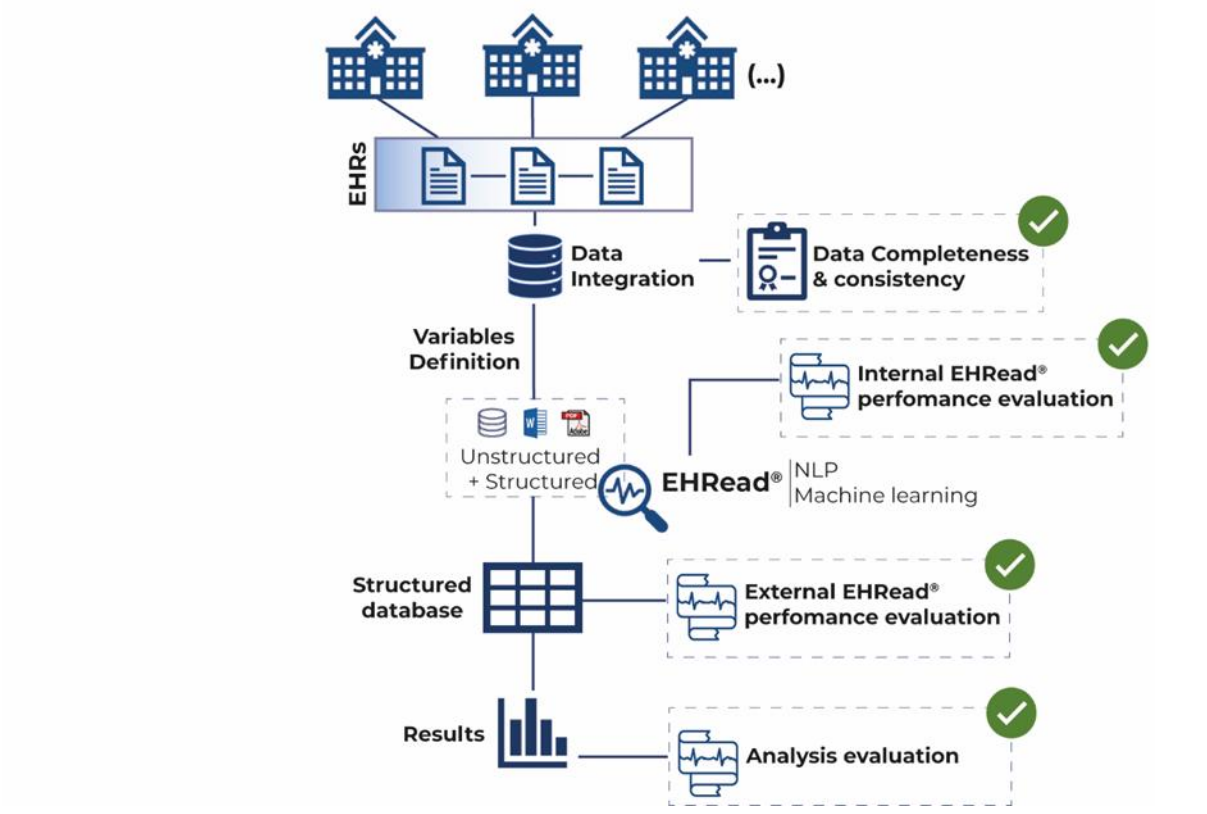

Figure 3. Overall summary of Medsavana and Savana Research methodology. *EHR data from participating hospitals is integrated by Medsavana and assessed for data completeness. Following the definition of the study variables, the unstructured free text from EHRs is extracted using the EHread© technology and internal and external performance evaluation is performed. The resulting study database is used to analyze the data and generate the study results.*

7 LIMITATIONS OF THE STUDY

The following limitations should be considered for the study:

- Data Scope: Only information contained within the Electronic Health Records (EHRs) can be analyzed. This inherently limits the study to the data that has been recorded and excludes any information that might not have been documented.
- Data Completeness: There might be missing or incomplete data across the different databases included in this study. Given that multiple databases from various hospitals will be incorporated, inconsistencies and gaps in medical data collection are likely to be encountered. This could affect the comprehensiveness and accuracy of the analysis.
- Standardization Issues: The lack of standardization in EHRs presents several challenges. Variations in the type of data collected across different disciplines, the use of standard versus in-house medical terminology, and the potential for omitted information or misuse of sections within the records are limitations that should be considered. These discrepancies can lead to difficulties in data integration and interpretation.
- NLP Limitations: The use of Natural Language Processing (NLP) to extract data from EHRs may introduce additional limitations. NLP algorithms may not accurately capture the context or nuances of medical language, leading to potential misinterpretations or errors in data extraction.
- Temporal Relevance: The data extracted from EHRs represents a snapshot in time and may not reflect changes in patient conditions or treatments over time. This temporal limitation could impact the study's ability to draw longitudinal conclusions.

8 ETHICAL CONSIDERATIONS

8.1 Patient information

The Study will be executed in compliance with the applicable law, in specific, with the Organic Law 3/2018 of December 5, 2018, on the Protection of Personal Data and guarantee of digital rights and in compliance with RD 957/2020 regulating observational studies with medicinal products for human use.

This Study will be executed with an anonymized Database. In order to execute the anonymization process, it is necessary to have access to the Participating Site's database, which requires a prior technical process of pseudonymization of the EHRs performed by the Participating Site.

The data anonymization process is divided into three stages:

8.1.1 *Submission process*

For the purposes of this section, EHRs shall be those that the Participating Site transfers to Medsavana according to the following: the structured data, data relating to the patient and episode ID, date of birth, EHR number shall be fully pseudonymized, in such a way as to not allow Medsavana to identify any of the patients. Given that the process of pseudonymization such personal data is performed by the Participating Site prior to the delivery of the personal data to Medsavana, the additional information that would allow the reversion of the data and, therefore, their identification or re-identification, will always appear separately, and Medsavana will not be able to access such information.
Thus, in accordance with the foregoing paragraph, the pseudonymization process must ensure the suppression of the patients' names and surnames, as well as replacing the patient ID (EHRs identification number used in the Participating Site) with a randomly generated identifier that does not correspond to the ID used by the Participating Site.
The Participating Site shall transfer the pseudonymized personal data to Medsavana through the means agreed between the Parties. The conversion table identifying each individual patient and containing the association between the patient ID used at the Participating Site and the new identification number under which the Participating Site will communicate the pseudonymized data to Medsavana shall be retained in the Participating Site's systems and infrastructure, without Medsavana having access to such conversion table. The Participating Site shall keep this conversion table for as long as necessary and in any case until the end of the Study, as well as for the periods of limitation of legal obligations as stated in the applicable regulations.

8.1.2 *Anonymization process*

Medsavana will apply NLP techniques to the pseudonymized data to find and extract the clinical variables described in the Study Protocol that will subsequently be structured, generating a database that only contains clinical terms (based on the SnomedCT ontology), and which therefore no longer

contains pseudonymized free text. On such database, which only contains variables in numeric format, anonymization techniques are applied to the following identifiers:

- The patient identifier (an internal system for sample naming which is not related with the patient or EHR number) called patient_client_id.
- The document identifier.
- The episode identifier.
- The hospital identifier.
- The date of birth.

These anonymization techniques that Medsavana will apply will consist of the following:

- For the date of birth, Medsavana will randomize the day of all records by the 15th day for each of them.
- For the rest of the identifiers described, i.e., patient identifier, patient-client identifier, document identifier, episode identifier, and hospital identifier, a one-way Hash function will be applied that converts the original values into a unique code with different digits.

Specifically, Medsavana uses a one-way Hash function as a "Hash with salt" method. This is a type of keyed hash algorithm that is generated from a specific hash function and is used as a hash-based message authentication code (HMAC). The HMAC process mixes a secret key with the message data, hashes the result with the hash function, re-mixes that hash value with the secret key, and applies the hash function a second time.

An HMAC can be used to determine whether a message sent over an insecure channel has been manipulated, provided that the sender and receiver share a secret key. The sender calculates the hash value of the original data and sends the original data and the hash value as a single message. The receiver recalculates the hash value in the received message and checks that the calculated HMAC matches the transmitted HMAC.

Any change in the data or in the hash value results in a mismatch, as it is necessary to know the secret key in order to change the message and reproduce the correct hash value. Therefore, if the original and the calculated hash value match, the message is authenticated.

The pseudonymised numeric identifiers are replaced by other identifiers through an irreversible process, so the result of the process is an anonymised database with no free text. This anonymised database is transferred from Medsavana to Savana Research which receives the aggregated and anonymised database.

8.1.3 *Transmission and verification of the database*

Once the anonymization process of the Participating Site's database has been completed, this, together with the databases of the rest of the Participating Sites in the Study, will constitute the final database of the Study. Only this anonymized and aggregated database will be transferred by Medsavana to Savana Research. In such a way, data will only be communicated to Savana Research once they have been anonymized and integrated into the final database of the Study, and Savana

Research will not have access to any pseudonymized data corresponding to the initial submission made by the Participating Site to Medsavana.

Once the database has been verified and confirmed by Savana Research, Medsavana will execute a process whereby the intermediate data stored in the Medsavana repositories will be blocked in order to make it impossible to reverse the anonymization process.

8.2 Patient´s consent exemption

The data to be used in this study will be obtained directly and retrospectively from the EHRs of the participating hospitals using Artificial Intelligence techniques. Data for the Study will be extracted by the IT departments of the participating hospitals. No data that could identify the patients directly will appear in this file as the patients´ ID will be presented in a pseudonymized format. This data also will be protected using the security measures described in **11.2**, **Annex II**.

In the performance of this study there will not be any data report form. The applied methodology involves a secondary use of the data contained in the EHR´s through the combination of existing structured and unstructured data, therefore, it does not involve the direct interaction with the study participant patients.

Regarding the need to collect the informed consent of the participants in the study, the exemption of obtaining such consent is submitted to the evaluation of the corresponding Ethics Committee by taking into consideration the following reasons:

1. The data is submitted to Medsavana by the hospital in pseudonymized format, and the additional information that would allow the data to be reversed will be kept in the hospital. Medsavana will not be able to access this information. Therefore, the patients could not be directly re-identified. See sections **8.1.1** and **8.1.2** for more details.
2. The study involves secondary use of the EHRs and therefore there would be no interaction with any of the participant patients. This study does not entail any risk for such participants.
3. The objectives to be achieved by conducting the observational study are of significant value in society in relation to the disease studied and the clinical breakthroughs it aims to demonstrate.
4. Lastly, obtaining the informed consent would be impossible considering that the data will be received in pseudonymized format, and it will imply a disproportionate effort due to the amount of EHRs that will be collected for the study.

Considering the above-mentioned reasons no patient´s informed consent will be required in this study and that is why the exemption of such consent is requested to the Ethics Committees, as applicable.

Institutional review board (IRB)/Independent ethics committee (IEC)

There must be prospective approval of the study protocol, protocol amendments, and other relevant documents from the IRBs/IECs. All correspondence with the IRB/IEC must be retained. Copies of IRB/IEC approvals must be forwarded to Sponsor.

Before the start of data collection, the study will be presented for review or notification to a national or central IRB or IEC in the designated country, as and if required by local regulations. Additionally,

the study will be presented or notified to regional and site IEC/IRBs, as and if required by local laws or regulations and/or hospital policies.

All amendments to the protocol will be subject to the appropriate review and approval process, in accordance with local regulations. At the end of the study, when required by local regulations, the participating physician or participating site director (or the funding company where necessary) will notify the IEC of the completion of the study.

Ethical conduct of the study

The study will be conducted in accordance with legal and regulatory requirements, as well as with scientific purpose, value, and rigor, and follow generally accepted research practices described in the International Society for Pharmacoepidemiology (ISPE) Guidelines for Good Pharmacoepidemiology Practices (GPP) and applicable regulatory requirements, the Helsinki Declaration in its latest edition, GPP, and applicable local regulations.

Adverse Event (AE), Pregnancy Exposure, and Incident Reporting

The use of automated methods for data extractions means there is no potential to collect serious and non-serious AEs, pregnancy exposures, or incidents related to any **Sponsor** product during the conduct of this research, as the minimum criteria of the identifiable patient, reporter, suspect product, and event, needed to collect and report individual case safety reports do not present in the data source.

This study involves a combination of existing structured data and unstructured data, which will be converted to structured form during the implementation of the protocol solely by a computer using automated/algorithmic methods, such as natural language processing.

In these data sources, individual patient data are not retrieved or validated, it is not possible to link (i.e., identify a potential association between) a particular product and medical event for any individual. Thus, the minimum criteria for reporting an adverse event (AE) (i.e., identifiable patient, identifiable reporter, a suspect product, and event) cannot be met.

Besides, the design of this study is characterized by secondary use of data previously collected from consumers or healthcare professionals for other purposes. For these studies, according to RD 957/2020, the submission of suspected adverse reactions in the form of ICSRs is not required.

9 STUDY MANAGEMENT

9.1 Database Retention and Archiving of Study Documents

The investigator must retain all study records and source documents for the maximum period required by applicable regulations and guidelines, institutional procedures, or for the period specified by the Sponsor, whichever is longest. The investigator must contact the Sponsor prior to destroying any records associated with the study.

9.2 Changes in the Protocol

Subsequent changes in protocol after its approval will be made only when major changes in the included population or the primary objectives are needed. These updates to the protocol will require an amendment and must be dated and signed. Amendments to the protocol should not be implemented without the prior approval of the IEC, when appropriate, or when the competent authorities have alleged reasons for not accepting them. Documentation of the approval of the amendments by the IEC should be provided to the Sponsor, if applicable. When the changes only affect the logistical or administrative aspects of the study, it is only necessary to inform the IEC.

9.3 Registration of Study on Public Website

The study will be registered in the following official registries:

Spanish Clinical Studies Registry (Registro Español de Estudios Clínicos, REec): https://reec.aemps.es/reec/public/web.html

HMA-EMA Catalogue of Real-World Data Studies: https://catalogues.ema.europa.eu/catalogue-rwd-studies

9.4 Publication Policy

In accordance with Good Publication Practices and the guidelines of the International Committee of Medical Journal Editors, the database owner/Sponsor will have the right to publish the main data and information (i.e., multicenter data) without the authorization of the participating physicians.

Participating physicians will recognize the integrity of a multicenter study by not sending for publication data from any of their hospitals until the combined results of the study are sent for publication once completed, within 12 months after the final data become available.

Authorship of publications derived from this study will be based on authorship standards, such as those described in the Uniform for Manuscripts Submitted to Biomedical Journals, which indicate that the cited authors must have made an important contribution to the design of the study or to the analysis and interpretation of the data, provided a critical review of the article and approved the final version.

Any work created in connection with the conduct of the study and content in the data that may benefit from the protection of copyright (except for the publications of the participating physicians) will be the property of the database owner and Sponsor as author and owner of the rights of said work.

Since the results of the study will be the sole and exclusive property of the Sponsor, the participating researchers and the study personnel acknowledge the irrevocable and complete assignment to Sponsor, before and during participation in the study, of their intellectual and industrial property rights over the results.

The publication of data, information, or results of the study by investigators must safeguard, in any case, the industrial and intellectual property rights of the Sponsor. In the event that Sponsor determines that the proposed publication could reveal secrets or business information owned by Sponsor, Savana will ensure that all participating researchers and other authors, at Sponsor's request, remove any such information from the publication prior to disclosure.

No participants in the study shall issue any press release or other form of publicity relating to this study without the prior written consent of the database owner/Sponsor. However, Sponsor may publish press releases that deal with content exclusively related to the study, its implementation, monitoring or results.

10 REFERENCES

1. Falanga A, Ay C, Di Nisio M, Gerotziafas G, Jara-Palomares L, Langer F, et al. Venous thromboembolism in cancer patients: ESMO Clinical Practice Guideline. Annals of Oncology. 2023;34(5):452-67.
2. Khorana AA, Francis CW, Culakova E, Kuderer NM, Lyman GH. Thromboembolism is a leading cause of death in cancer patients receiving outpatient chemotherapy. J Thromb Haemost. 2007;5(3):632-4.
3. Falanga A, Schieppati F, Russo L. Pathophysiology 1. Mechanisms of Thrombosis in Cancer Patients. Cancer Treat Res. 2019;179:11-36.
4. Mulder FI, Horvath-Puho E, van Es N, van Laarhoven HWM, Pedersen L, Moik F, et al. Venous thromboembolism in cancer patients: a population-based cohort study. Blood. 2021;137(14):1959-69.
5. Khorana AA, Connolly GC. Assessing risk of venous thromboembolism in the patient with cancer. J Clin Oncol. 2009;27(29):4839-47.
6. Blom JW, Doggen CJ, Osanto S, Rosendaal FR. Malignancies, prothrombotic mutations, and the risk of venous thrombosis. JAMA. 2005;293(6):715-22.
7. Falanga A. The incidence and risk of venous thromboembolism associated with cancer and nonsurgical cancer treatment. Cancer Invest. 2009;27(1):105-15.
8. Nalluri SR, Chu D, Keresztes R, Zhu X, Wu S. Risk of venous thromboembolism with the angiogenesis inhibitor bevacizumab in cancer patients: a meta-analysis. JAMA. 2008;300(19):2277-85.
9. Starling N, Rao S, Cunningham D, Iveson T, Nicolson M, Coxon F, et al. Thromboembolism in patients with advanced gastroesophageal cancer treated with anthracycline, platinum, and fluoropyrimidine combination chemotherapy: a report from the UK National Cancer Research Institute Upper Gastrointestinal Clinical Studies Group. J Clin Oncol. 2009;27(23):3786-93.
10. Wang TF, Carrier M. Immune Checkpoint Inhibitors-Associated Thrombosis: Incidence, Risk Factors and Management. Curr Oncol. 2023;30(3):3032-46.
11. Horsted F, West J, Grainge MJ. Risk of venous thromboembolism in patients with cancer: a systematic review and meta-analysis. PLoS Med. 2012;9(7):e1001275.
12. Munoz AJ, Souto JC, Lecumberri R, Obispo B, Sanchez A, Aparicio J, et al. Development of a predictive model of venous thromboembolism recurrence in anticoagulated cancer patients using machine learning. Thromb Res. 2023;228:181-8.
13. Lyman GH, Carrier M, Ay C, Di Nisio M, Hicks LK, Khorana AA, et al. American Society of Hematology 2021 guidelines for management of venous thromboembolism: prevention and treatment in patients with cancer. Blood Advances. 2021;5(4):927-74.
14. Streiff MB, Milentijevic D, McCrae K, Yannicelli D, Fortier J, Nelson WW, et al. Effectiveness and safety of anticoagulants for the treatment of venous thromboembolism in patients with cancer. Am J Hematol. 2018;93(5):664-71.
15. Patel HK, Khorana AA. Anticoagulation in Cancer Patients: a Summary of Pitfalls to Avoid. Curr Oncol Rep. 2019;21(2):18.
16. Carmona-Bayonas A, Gómez D, Martínez De Castro E, Pérez Segura P, Muñoz Langa J, Jimenez-Fonseca P, et al. A snapshot of cancer-associated thromboembolic disease in 2018–2019: First data from the TESEO prospective registry. European Journal of Internal Medicine. 2020;78:41-9.
17. Pina E, Antonio M, Peris J, Rossello E, Domenech P, Penafiel J, et al. Bemiparin as a long-term treatment for venous thrombosis in cancer patients: the ELEBAMA study. Clin Transl Oncol. 2020;22(4):616-20.
18. Ozaslan E, Ozkan M, Cicin I, Benekli M, Kocer M, Uysal M, et al. Effectiveness and Safety of LMWH Treatment in Patients With Cancer Diagnosed With Non-High-Risk Venous Thromboembolism:

Turkish Observational Study (TREBECA). Clinical and Applied Thrombosis/Hemostasis. 2018;24(6):973-9.
19. Galeano-Valle F, Pérez-Rus G, Demelo-Rodríguez P, Ordieres-Ortega L, Ortega-Morán L, Muñoz-Martín AJ, et al. Monitoring anti-Xa levels in patients with cancer-associated venous thromboembolism treated with bemiparin. Clinical and Translational Oncology. 2020;22(8):1312-20.
20. Muñoz Martín AJ, Palacios ML, Souto JC, Obispo B, Aparicio J, Garcia-Palomo A, et al. Predicting major bleeding events in anticoagulated cancer patients with venous thromboembolism using real-world data and machine learning. ASCO Annual Meeting: Journal of Clinical Oncology; 2022. p. suppl.e18744.
21. Izquierdo JL, Soriano JB, Gonzalez Y, Lumbreras S, Ancochea J, Echeverry C, et al. Use of N-Acetylcysteine at high doses as an oral treatment for patients hospitalized with COVID-19. Sci Prog. 2022;105(1):368504221074574.
22. Canales L, Menke S, Marchesseau S, D'Agostino A, Del Rio-Bermudez C, Taberna M, et al. Assessing the Performance of Clinical Natural Language Processing Systems: Development of an Evaluation Methodology. JMIR Med Inform. 2021;9(7):e20492.
23. Arguello-Gonzalez G, Aquino-Esperanza J, Salvador D, Breton-Romero R, Del Rio-Bermudez C, Tello J, et al. Negation recognition in clinical natural language processing using a combination of the NegEx algorithm and a convolutional neural network. BMC Med Inform Decis Mak. 2023;23(1):216.
24. Hernández Medrano I, Tello Guijarro J, Belda C, Urena A, Salcedo I, Espinosa-Anke L, et al. Savana: Re-using Electronic Health Records with Artificial Intelligence. Int J Interact Multimed Artif Intell. 2017;4:8-12.
25. Loscertales J, Abrisqueta-Costa P, Gutierrez A, Hernandez-Rivas JA, Andreu-Lapiedra R, Mora A, et al. Real-World Evidence on the Clinical Characteristics and Management of Patients with Chronic Lymphocytic Leukemia in Spain Using Natural Language Processing: The SRealCLL Study. Cancers (Basel). 2023;15(16).
26. Gonzalez-Juanatey C, Anguita-Sanchez M, Barrios V, Nunez-Gil I, Gomez-Doblas JJ, Garcia-Moll X, et al. Impact of Advanced Age on the Incidence of Major Adverse Cardiovascular Events in Patients with Type 2 Diabetes Mellitus and Stable Coronary Artery Disease in a Real-World Setting in Spain. J Clin Med. 2023;12(16).
27. Ancochea J, Izquierdo JL, Soriano JB. Evidence of Gender Differences in the Diagnosis and Management of Coronavirus Disease 2019 Patients: An Analysis of Electronic Health Records Using Natural Language Processing and Machine Learning. J Womens Health (Larchmt). 2021;30(3):393-404.
28. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.
29. Trujillo-Santos J, Farge-Bancel D, Pedrajas JM, Gomez-Cuervo C, Ballaz A, Braester A, et al. Enoxaparin versus dalteparin or tinzaparin in patients with cancer and venous thromboembolism: The RIETECAT study. Res Pract Thromb Haemost. 2022;6(4):e12736.
30. Guideline on clinical investigation of medicinal products for the treatment of venous thromboembolic disease. European Medicines Agency (EMA); 2016.
31. Schulman S, Kearon C. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. Journal of Thrombosis and Haemostasis. 2005;3(4):692-4.
32. Kaatz S, Ahmad D, Spyropoulos AC, Schulman S. Definition of clinically relevant non-major bleeding in studies of anticoagulants in atrial fibrillation and venous thromboembolic disease in non-surgical patients: communication from the SSC of the ISTH. Journal of Thrombosis and Haemostasis. 2015;13(11):2119-26.
33. Gonzalez-Juanatey C, Anguita-Sa Nchez M, Barrios V, Nunez-Gil I, Gomez-Doblas JJ, Garcia-Moll X, et al. Assessment of medical management in Coronary Type 2 Diabetic patients with previous

percutaneous coronary intervention in Spain: A retrospective analysis of electronic health records using Natural Language Processing. PLoS One. 2022;17(2):e0263277.
34. Graziani D, Soriano JB, Del Rio-Bermudez C, Morena D, Diaz T, Castillo M, et al. Characteristics and Prognosis of COVID-19 in Patients with COPD. J Clin Med. 2020;9(10).
35. Izquierdo JL, Ancochea J, Savana C-RG, Soriano JB. Clinical Characteristics and Prognostic Factors for Intensive Care Unit Admission of Patients With COVID-19: Retrospective Study Using Machine Learning and Natural Language Processing. J Med Internet Res. 2020;22(10):e21801.
36. Izquierdo JL, Almonacid C, Gonzalez Y, Del Rio-Bermudez C, Ancochea J, Cardenas R, et al. The impact of COVID-19 on patients with asthma. Eur Respir J. 2021;57(3).
37. Montoto C, Gisbert JP, Guerra I, Plaza R, Pajares Villarroya R, Moreno Almazan L, et al. Evaluation of Natural Language Processing for the Identification of Crohn Disease-Related Variables in Spanish Electronic Health Records: A Validation Study for the PREMONITION-CD Project. JMIR Med Inform. 2022;10(2):e30345.
38. Gomollon F, Gisbert JP, Guerra I, Plaza R, Pajares Villarroya R, Moreno Almazan L, et al. Clinical characteristics and prognostic factors for Crohn's disease relapses using natural language processing and machine learning: a pilot study. Eur J Gastroenterol Hepatol. 2022;34(4):389-97.
39. Espinosa L, Tello J, Pardo A, Medrano I, Ureña A, Salcedo I, et al. Savana: A Global Information Extraction and Terminology Expansion Framework in the Medical Domain. Procesamiento del Lenguaje Natural. 2016;57:23-30.
40. Benson T. Principles of Health Interoperability HL7 and SNOMED: Springer; 2nd edition; 2012.
41. Van Rijsbergen CJ. Information Retrieval - 2nd edition: Butterworth-Heinemann; 1979.
42. Ford E, Carroll JA, Smith HE, Scott D, Cassell JA. Extracting information from the text of electronic medical records to improve case detection: a systematic review. J Am Med Inform Assoc. 2016;23(5):1007-15.
43. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276-82.
44. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform. 2014;48:193-204.
45. Lu MJ, Zhong WH, Liu YX, Miao HZ, Li YC, Ji MH. Sample Size for Assessing Agreement between Two Methods of Measurement by Bland-Altman Method. Int J Biostat. 2016;12(2).
46. Kradjian S, Gutheil J, Baratelle AM, Einstein SG, Kaslow DC. Development of a Charter for an Endpoint Assessment and Adjudication Committee. Ther Innov Regul Sci. 2005;39:53-61.
47. Bemiparin: Summary of Product Characteristics. Spanish Agency of Medicines and Medical Devices.
48. Dalteparin: Summary of Product Characteristics. Spanish Agency of Medicines and Medical Devices.
49. Enoxaparin: Summary of Product Characteristics. Spanish Agency of Medicines and Medical Devices.
50. Prescribing Prophylactic and Therapeutic Doses of Enoxaparin Including Dose Banding Clinical Guideline. NHS Trust; 2024.
51. Tinzaparin: Summary of Product Characteristics. Spanish Agency of Medicines and Medical Devices.

11 ANNEXES

11.1 Annex I. Dosing Regimens for Bemiparin, Dalteparin, Enoxaparin, and Tinzaparin

Supplemental Table 1. Dosing regimens for Bemiparin, Dalteparin, Enoxaparin, and Tinzaparin

LMWH	Dosage	IU
Bemiparin (47)	Weight-Based Dosage	▪ 115 IU/kg ± 20% once daily
	Every 24 h Fixed Dosage (Based on weight)	▪ <50 kg: 5000 IU anti-Xa ▪ 50–70 kg: 7500 IU anti-Xa ▪ 70–100 kg: 10,000 IU anti-Xa ▪ 100–120 kg: 12,500 IU anti-Xa ▪ 120 kg: Adjust dose to 115 IU anti-Xa/kg/day
Dalteparin (48)	Weight-Based Dosage	▪ 200 IU/kg ± 20% once daily during the first month, followed by 150 IU/kg ± 20% once daily ▪ 100 IU/kg ± 20% twice daily continuously
	Every 24 h Fixed Dosage (Based on weight)	▪ <56 kg: 7500 IU ▪ 57–68 kg: 10,000 IU ▪ 69–82 kg: 12,500 IU ▪ 83–98 kg: 15,000 IU ▪ ≥99 kg: 18,000 IU
Enoxaparin (49, 50)	Weight-Based Dosage:	▪ 100 IU/kg ± 20% twice daily ▪ 150 IU/kg ± 20% once daily
	Every 24 h Fixed Dosage (Based on weight)	▪ 40-43.9 kg: 6000 IU ▪ 44-50.9 kg: 7000 IU ▪ 51-56.9 kg: 8000 IU ▪ 57-63.9 kg: 9000 IU ▪ 64-66.9 kg: 10,000 IU ▪ 67-73.9 kg: 10,500 UI ▪ 74-84.9 kg: 12,000 UI ▪ 85-94.9 kg: 13,500 UI ▪ 95-103.9 kg: 15,000 UI ▪ 104-109.9 kg: 16,000 UI ▪ 110-116.9 kg: 17,000 UI ▪ 117-120 kg: 18,000 UI
	Every 12 h Fixed Dosage (Based on weight)	▪ 45–54 kg: 5000 IU ▪ 55–64 kg: 6000 IU ▪ 65–74 kg: 7000 IU ▪ 75–84 kg: 8000 IU ▪ 85–94 kg: 9000 IU ▪ 95–104 kg: 10,000 IU ▪ 105-109 Kg: 11,000 UI ▪ 110-120 Kg: 12,000 UI
Tinzaparin (51)	Weight-Based Dosage	▪ 175 IU/kg ± 20% once daily
	Every 24 h Fixed Dosage (Based on weight)	▪ 32-37 kg: 6000 IU ▪ 38-42 kg: 7000 IU

		▪ 43-48 kg: 8000 IU ▪ 49-54 kg: 9000 IU ▪ 55-59 kg: 10,000 IU ▪ 60-65 kg: 11,000 IU ▪ 66-71 kg: 12,000 IU ▪ 72-77 kg: 13,000 IU ▪ 78-82 kg: 14,000 IU ▪ 83-88 kg: 15,000 IU ▪ 89-94 kg: 16,000 IU ▪ 95-99 kg: 17,000 IU ▪ 100-105 kg: 18,000 ▪ <32 kg or >105 kg: Adjust dose to 175 IU/kg/day

11.2 Annex II. Security measures

This section sets out the technical and organizational security measures to be implemented by Medsavana in the performance of data processing activities involved in the Study.

a. Logical access control to the systems processing personal data

As a minimum, the following security measures shall be implemented:

1. Access control:
    - Access to systems subject to authentication of authorized personnel.
    - The existence of an updated list of users with authorized access to the information systems.
    - Prohibition of the use of anonymous or generic accounts, except in justified and limited situations.
    - Implementation of an access management system. Access administration should be centralized, and access authorization should only be granted to authorized personnel, who shall be the only personnel that can grant, alter, or revoke access to systems.
    - Segregation of duties in IT systems, which prevents a single individual from accessing, modifying, or using assets without authorization or detection.
    - If outside personnel have access to the resources, they shall be subject to the same security conditions and obligations as in-house personnel.
2. Identification and authentication:
    - Use of passwords with specific security parameters (uppercase, lowercase, numbers, letters and special characters, minimum number of 6 characters, and expiration once a year), and keeping them unintelligible.
    - Use of a procedure for assigning, distributing, and storing passwords that ensures their confidentiality and integrity.
3. Media management:
    - Identification and inventory of the devices processing personal data, as well as of the users accessing them.
    - Adoption of measures aimed at preventing subsequent access or recovery of the information contained in the media once it has been decided to dispose of them. To this end, they must be destroyed or completely erased utilizing secure erasure systems. Devices containing personal data should be physically destroyed or the information should be destroyed, deleted or overwritten using techniques that do not allow for the recovery of the original information, rather than using normal erasure or formatting.
    - The distribution of media containing personal data that pose a high risk to the rights and freedoms of data subjects shall be carried out after encrypting such data or using any other mechanism that ensures that such information is not accessible or manipulated during transport.
4. Back-up and continuity of service:

- Documentation of back-up and recovery procedures, which ensure that they are at all times reconstructed to the state they were in at the time of loss or destruction.
- Periodic incremental and daily backups.
- Ensure that systems are operational and that failures are properly reported. Accurate and complete records of backups should be kept.
- The backups should be stored at a remote site, at a sufficient distance to be spared any damage from a disaster at the main site.
- The controls applied to the media at the main site should be extended to the location of the backup copies.
- A documented and tested Continuity Plan is in place that ensures production, data and services according to the agreed contract/SLA deadlines, and guarantees at least:
    - That the electricity supply installations are adequate for the equipment they are intended to support.
    - That emergency disconnectors are installed near the emergency exits of the rooms where the equipment is located to facilitate quick disconnection in case of an emergency.
    - That, if the main power supply fails, emergency lighting is available.

5. Developmental testing
    - The most significant changes are identified and recorded.
    - Changes are planned and tested before implementation.
    - The risks and potential impacts of such changes are assessed and subject to a formal approval process.
    - Users will have to use different profiles for production and test systems.
    - No pre-implementation testing or modification of information systems with personal data shall be carried out with real data. Decoupled data shall be used.
    - If deemed necessary, the same security measures must be implemented as in the production environment, and a backup copy must be made beforehand.
    - Segregation of IT test and production environments.
6. Network security controls
    - Use of firewall, router, and VPN-based access controls to protect private service networks and back-end servers.
    - Infrastructure security with ad-hoc monitoring.
    - Logging of access to host servers, applications, databases, routers, switches, etc.
    - The transmission of personal data over public or wireless electronic communications networks shall be carried out under secure protocols.
    - Sensitive personal data will be encrypted during transmission using security protocols utilizing strong algorithms and encryption keys.
    - There is a limited existence of system administrators.

- Adequate measures are in place to log system administrators' access to the infrastructure.
- Establish mechanisms for recording actions on personal data or logging.

b. Physical access control to facilities and personal data processing areas

At least some of the following security measures have been implemented to prevent physical access to workplaces and data processing centers:

- Access control system.
- Identity reader, magnetic card, or chip card.
- Provision of keys.
- Lock on the door (automatic doors, etc.).
- Alarm system, video, and video surveillance monitors.
- Register of entries and exits from the facilities.
- Location of data centers in secure facilities that provide physical security, redundant power, and infrastructure redundancy.

c. Register of incidents

The following security measures have been implemented:

- Existence of a procedure for notification and management of incidents affecting personal data.
- Existence of a procedure for notification and management of security breaches or violations, and their notification in due time and form to the Data Controller, to comply with the requirements of the regulations.
- Maintenance of a record of incidents/violations.
- Training actions are carried out for employees and other people with access to data, to guarantee adequate data processing under the requirements of the regulations.
- A process of regular verification, evaluation, and assessment of the effectiveness of the technical and organizational measures to ensure the security of the processing has been established, taking into account in particular the risks related to the data processing.

Certificate Of Completion

Envelope Id: 67907A5E-CDC8-4EC5-BAC2-6EFF4B5E1CE0
Subject: ROVI_BEMICAT_Protocol_V1.2_11JUN2025
Source Envelope:
Document Pages: 50
Certificate Pages: 8
AutoNav: Enabled
EnvelopeId Stamping: Disabled
Time Zone: (UTC+01:00) Brussels, Copenhagen, Madrid, Paris

Signatures: 6
Initials: 0

Status: Completed

Envelope Originator:
Gema Jara Nieto
JULIAN CAMARILLO, 35
Tax-ID: A28041283
MADRID, Madrid 28037
gjara@rovi.es
IP Address: 213.99.29.109

Record Tracking

Status: Original
17-06-2025 | 08:55

Holder: Gema Jara Nieto
gjara@rovi.es

Location: DocuSign

Signer Events	Signature	Timestamp
Ana Belén López Yélamos ablopez@rovi.es EU-QPPV LABORATORIOS FARMACEUTICOS ROVI, S.A. Security Level: Email, Account Authentication (Required)	 Signature Adoption: Pre-selected Style Signature ID: 02503044-FC15-44E4-B1A5-D32A7AA01F4A Using IP Address: 213.99.29.109 With Signing Authentication via Docusign password With Signing Reasons (on each tab): I approve this document	Sent: 17-06-2025 | 09:12 Resent: 17-06-2025 | 11:47 Viewed: 17-06-2025 | 12:17 Signed: 17-06-2025 | 12:47
Electronic Record and Signature Disclosure: Accepted: 22-11-2022 | 14:44 ID: 6394a6f3-a6ab-4e39-8d84-6fb5a2a8f484		
Carlos Arias carias@savanamed.com Security Level: Email, Account Authentication (Required)	 Signature Adoption: Uploaded Signature Image Signature ID: 065D0B91-AA7E-4A56-BE8C-AA860EEDAC7D Using IP Address: 149.74.97.39 With Signing Authentication via Docusign password With Signing Reasons (on each tab): Apruebo este documento	Sent: 17-06-2025 | 09:12 Viewed: 17-06-2025 | 10:54 Signed: 17-06-2025 | 10:56
Electronic Record and Signature Disclosure: Accepted: 28-02-2025 | 10:24 ID: df591307-9d9b-47b9-b37f-5a6febc392e1		

Signer Events	Signature	Timestamp
Carlos Salazar García csalazar@rovi.es Clinical Development Physician ROVI Security Level: Email, Account Authentication (Required) **Electronic Record and Signature Disclosure:** Accepted: 08-01-2025 | 13:59 ID: 9445881f-bad6-4a43-82d8-14d699af54ab	Carlos Salazar García Signature Adoption: Pre-selected Style Signature ID: E350286C-01EC-46D3-9E97-AB655F07AB02 Using IP Address: 213.99.29.109 With Signing Authentication via Docusign password With Signing Reasons (on each tab): I approve this document	Sent: 17-06-2025 | 09:12 Viewed: 17-06-2025 | 09:24 Signed: 17-06-2025 | 09:26
David Casadevall dcasadevall@savanamed.com Security Level: Email, Account Authentication (Required) **Electronic Record and Signature Disclosure:** Accepted: 28-02-2025 | 13:03 ID: 1d00bf11-28c9-4d9c-b9c8-745897c3c964	David Casadevall Signature Adoption: Pre-selected Style Signature ID: 0EC07954-E7C4-49AF-88EA-CB8C57D8D473 Using IP Address: 31.4.247.226 With Signing Authentication via Docusign password With Signing Reasons (on each tab): Apruebo este documento	Sent: 17-06-2025 | 09:12 Resent: 17-06-2025 | 11:47 Resent: 17-06-2025 | 15:33 Viewed: 17-06-2025 | 18:07 Signed: 17-06-2025 | 18:08
Eduard Sarró esarro@savanamed.com Security Level: Email, Account Authentication (Required) **Electronic Record and Signature Disclosure:** Accepted: 17-06-2025 | 18:00 ID: 2dbbf642-ec69-40e7-bf1a-0f7066abc09a	 Signature Adoption: Uploaded Signature Image Signature ID: 9291C6CF-F0A9-406D-8192-D902C2DDC32E Using IP Address: 37.223.196.138 With Signing Authentication via Docusign password With Signing Reasons (on each tab): I approve this document	Sent: 17-06-2025 | 09:12 Resent: 17-06-2025 | 11:47 Resent: 17-06-2025 | 15:33 Viewed: 17-06-2025 | 18:00 Signed: 17-06-2025 | 18:01
Natalia Polo npolo@savanamed.com Security Level: Email, Account Authentication (Required)	Natalia Polo Signature Adoption: Pre-selected Style Signature ID: EB4125D7-8843-4540-85C3-0EA48F77FA11 Using IP Address: 83.56.178.111 With Signing Authentication via Docusign password With Signing Reasons (on each tab): Apruebo este documento	Sent: 17-06-2025 | 09:12 Viewed: 17-06-2025 | 09:25 Signed: 17-06-2025 | 09:26

Signer Events	Signature	Timestamp
Electronic Record and Signature Disclosure: Accepted: 04-03-2025 | 09:07 ID: 955b6da1-7b74-480e-9d6b-6d6fc22c6e9d		

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	17-06-2025 | 09:12
Certified Delivered	Security Checked	17-06-2025 | 09:25
Signing Complete	Security Checked	17-06-2025 | 09:26
Completed	Security Checked	17-06-2025 | 18:08

Payment Events	Status	Timestamps

Electronic Record and Signature Disclosure

ELECTRONIC RECORD AND SIGNATURE DISCLOSURE

From time to time, Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. (we, us or Company) may be required by law to provide to you certain written notices or disclosures. Described below are the terms and conditions for providing to you such notices and disclosures electronically through the DocuSign system. Please read the information below carefully and thoroughly, and if you can access this information electronically to your satisfaction and agree to this Electronic Record and Signature Disclosure (ERSD), please confirm your agreement by selecting the check-box next to 'I agree to use electronic records and signatures' before clicking 'CONTINUE' within the DocuSign system.

Getting paper copies

At any time, you may request from us a paper copy of any record provided or made available electronically to you by us. You will have the ability to download and print documents we send to you through the DocuSign system during and immediately after the signing session and, if you elect to create a DocuSign account, you may access the documents for a limited period of time (usually 30 days) after such documents are first sent to you. After such time, if you wish for us to send you paper copies of any such documents from our office to you, you will be charged a $0.00 per-page fee. You may request delivery of such paper copies from us by following the procedure described below.

Withdrawing your consent

If you decide to receive notices and disclosures from us electronically, you may at any time change your mind and tell us that thereafter you want to receive required notices and disclosures only in paper format. How you must inform us of your decision to receive future notices and disclosure in paper format and withdraw your consent to receive notices and disclosures electronically is described below.

Consequences of changing your mind

If you elect to receive required notices and disclosures only in paper format, it will slow the speed at which we can complete certain steps in transactions with you and delivering services to you because we will need first to send the required notices or disclosures to you in paper format, and then wait until we receive back from you your acknowledgment of your receipt of such paper notices or disclosures. Further, you will no longer be able to use the DocuSign system to receive required notices and consents electronically from us or to sign electronically documents from us.

All notices and disclosures will be sent to you electronically

Unless you tell us otherwise in accordance with the procedures described herein, we will provide electronically to you through the DocuSign system all required notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you during the course of our relationship with you. To reduce the chance of you inadvertently not receiving any notice or disclosure, we prefer to provide all of the required notices and disclosures to you by the same method and to the same address that you have given us. Thus, you can receive all the disclosures and notices electronically or in paper format through the paper mail delivery system. If you do not agree with this process, please let us know as described below. Please also see the paragraph immediately above that describes the consequences of your electing not to receive delivery of the notices and disclosures electronically from us.

How to contact Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. :

You may contact us to let us know of your changes as to how we may contact you electronically, to request paper copies of certain information from us, and to withdraw your prior consent to receive notices and disclosures electronically as follows:
To contact us by email send messages to: ckrottenthaler@rovi.com

To advise Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. of your new email address

To let us know of a change in your email address where we should send notices and disclosures electronically to you, you must send an email message to us at ckrottenthaler@rovi.com and in the body of such request you must state: your previous email address, your new email address. We do not require any other information from you to change your email address.

If you created a DocuSign account, you may update it with your new email address through your account preferences.

To request paper copies from Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A.

To request delivery from us of paper copies of the notices and disclosures previously provided by us to you electronically, you must send us an email to ckrottenthaler@rovi.com and in the body of such request you must state your email address, full name, mailing address, and telephone number. We will bill you for any fees at that time, if any.

To withdraw your consent with Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A.

To inform us that you no longer wish to receive future notices and disclosures in electronic format you may:

i. decline to sign a document from within your signing session, and on the subsequent page, select the check-box indicating you wish to withdraw your consent, or you may;

ii. send us an email to ckrottenthaler@rovi.com and in the body of such request you must state your email, full name, mailing address, and telephone number. We do not need any other information from you to withdraw consent.. The consequences of your withdrawing consent for online documents will be that transactions may take a longer time to process..

Required hardware and software

The minimum system requirements for using the DocuSign system may change over time. The current system requirements are found here: https://support.docusign.com/guides/signer-guide-signing-system-requirements.

Acknowledging your access and consent to receive and sign documents electronically

To confirm to us that you can access this information electronically, which will be similar to other electronic notices and disclosures that we will provide to you, please confirm that you have read this ERSD, and (i) that you are able to print on paper or electronically save this ERSD for your future reference and access; or (ii) that you are able to email this ERSD to an email address where you will be able to print on paper or save it for your future reference and access. Further, if you consent to receiving notices and disclosures exclusively in electronic format as described herein, then select the check-box next to 'I agree to use electronic records and signatures' before clicking 'CONTINUE' within the DocuSign system.

By selecting the check-box next to 'I agree to use electronic records and signatures', you confirm that:

- You can access and read this Electronic Record and Signature Disclosure; and
- You can print on paper this Electronic Record and Signature Disclosure, or save or send this Electronic Record and Disclosure to a location where you can print it, for future reference and access; and
- Until or unless you notify Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. as described above, you consent to receive exclusively through electronic means all notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you by Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. during the course of your relationship with Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. .

ELECTRONIC RECORD AND SIGNATURE DISCLOSURE

From time to time, Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. (we, us or Company) may be required by law to provide to you certain written notices or disclosures. Described below are the terms and conditions for providing to you such notices and disclosures electronically through the DocuSign system. Please read the information below carefully and thoroughly, and if you can access this information electronically to your satisfaction and agree to this Electronic Record and Signature Disclosure (ERSD), please confirm your agreement by selecting the check-box next to 'I agree to use electronic records and signatures' before clicking 'CONTINUE' within the DocuSign system.

Getting paper copies

At any time, you may request from us a paper copy of any record provided or made available electronically to you by us. You will have the ability to download and print documents we send to you through the DocuSign system during and immediately after the signing session and, if you elect to create a DocuSign account, you may access the documents for a limited period of time (usually 30 days) after such documents are first sent to you. After such time, if you wish for us to send you paper copies of any such documents from our office to you, you will be charged a $0.00 per-page fee. You may request delivery of such paper copies from us by following the procedure described below.

Withdrawing your consent

If you decide to receive notices and disclosures from us electronically, you may at any time change your mind and tell us that thereafter you want to receive required notices and disclosures only in paper format. How you must inform us of your decision to receive future notices and disclosure in paper format and withdraw your consent to receive notices and disclosures electronically is described below.

Consequences of changing your mind

If you elect to receive required notices and disclosures only in paper format, it will slow the speed at which we can complete certain steps in transactions with you and delivering services to you because we will need first to send the required notices or disclosures to you in paper format, and then wait until we receive back from you your acknowledgment of your receipt of such paper notices or disclosures. Further, you will no longer be able to use the DocuSign system to receive required notices and consents electronically from us or to sign electronically documents from us.

All notices and disclosures will be sent to you electronically

Unless you tell us otherwise in accordance with the procedures described herein, we will provide electronically to you through the DocuSign system all required notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you during the course of our relationship with you. To reduce the chance of you inadvertently not receiving any notice or disclosure, we prefer to provide all of the required notices and disclosures to you by the same method and to the same address that you have given us. Thus, you can receive all the disclosures and notices electronically or in paper format through the paper mail delivery system. If you do not agree with this process, please let us know as described below. Please also see the paragraph immediately above that describes the consequences of your electing not to receive delivery of the notices and disclosures electronically from us.

How to contact Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. :

You may contact us to let us know of your changes as to how we may contact you electronically, to request paper copies of certain information from us, and to withdraw your prior consent to receive notices and disclosures electronically as follows:

To advise Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. of your new email address

To let us know of a change in your email address where we should send notices and disclosures electronically to you, you must send an email message to us at and in the body of such request you must state: your previous email address, your new email address.

If you created a DocuSign account, you may update it with your new email address through your account preferences.

To request paper copies from Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A.

To request delivery from us of paper copies of the notices and disclosures previously provided by us to you electronically, you must send us an email to and in the body of such request you must state your email address, full name, mailing address, and telephone number.

To withdraw your consent with Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A.

To inform us that you no longer wish to receive future notices and disclosures in electronic format you may:

i. decline to sign a document from within your signing session, and on the subsequent page, select the check-box indicating you wish to withdraw your consent, or you may;

ii. send us an email to and in the body of such request you must state your email, full name, mailing address, and telephone number. . .

Required hardware and software

The minimum system requirements for using the DocuSign system may change over time. The current system requirements are found here: https://support.docusign.com/guides/signer-guide-signing-system-requirements.

Acknowledging your access and consent to receive and sign documents electronically

To confirm to us that you can access this information electronically, which will be similar to other electronic notices and disclosures that we will provide to you, please confirm that you have read this ERSD, and (i) that you are able to print on paper or electronically save this ERSD for your future reference and access; or (ii) that you are able to email this ERSD to an email address where you will be able to print on paper or save it for your future reference and access. Further, if you consent to receiving notices and disclosures exclusively in electronic format as described herein, then select the check-box next to ‘I agree to use electronic records and signatures’ before clicking ‘CONTINUE’ within the DocuSign system.

By selecting the check-box next to ‘I agree to use electronic records and signatures’, you confirm that:

- You can access and read this Electronic Record and Signature Disclosure; and
- You can print on paper this Electronic Record and Signature Disclosure, or save or send this Electronic Record and Disclosure to a location where you can print it, for future reference and access; and
- Until or unless you notify Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. as described above, you consent to receive exclusively through electronic means all notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you by Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. during the course of your relationship with Part11 LABORATORIOS FARMACEUTICOS ROVI, S.A. .