DATA EXTRACTION

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.
Version:	1.1
Date:	27FEB2025
Study type:	Non interventional study

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.

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

Data Processing

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.

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.