D1.2 – ETHICS GUIDANCE REPORT, ETHICS IMPACT ASSESSMENT

Project Full Title: *AcceleRating the Translation of virtual twins towards a pErsonalised Management of steatotic liver patients*
Project acronym: *ARTEMIs*
Project type: Horizon Europe | RIA (Topic HORIZON-HLTH-2023-TOOL-05-03)
Grant agreement no: 101136299

Document information:

Deliverable no.	**D1.2**
Title	Ethics guidance report, Ethics Impact Assessment
Work Package	9
Dissemination level	Public
Nature	Report
Responsible partner	MAT and MEDEX-BCP
Main Contact person	Amelia Suarez (MAT) and Laure Saint-Aubert (MEDEX)
Deliverable due submission date	30/09/2024
Deliverable actual submission date	09/10/2024

Document validation:

Validated by	Name	Organisation short name	Visa
Responsible	Laura Muñoz Manon Pelissier	MAT MEDEX-BCP	OK OK
Reviewer 1	N/A	VHIR	N/A
Reviewer 2	Florian Leuschner	UKHD	OK

Disclaimer

The opinions stated in this report reflect the opinion of the authors and not the opinion of the European Commission.

History of changes:

Change explanation	Pages affected	Change made by (Name & surname)	Date

Table of Contents

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

Table of Figures

Table of Tables

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

INTRODUCTION

The overall objective of the ARTEMIs project is to develop and demonstrate the feasibility of a clinical decision support system (CDSS) for use in the clinical management of metabolic steatotic liver disease (MASLD is the new nomenclature, previously MASLD). The CDSS aims to provide clinicians with clinically meaningful information for more personalised management of MASLD patients. Such developments require representative, multi-domain and multi-centre patient data. The ARTEMIs cohort is a key component of the project, with eleven participating clinical research centres from France, Spain, Germany, Belgium, Austria, Italy and the UK.

This deliverable aims to define the project's ethical framework and ensure ethical compliance in conducting the cohort study, using personal data in the research activities, and designing and implementing an ethical ARTEMIs CDSS.

This report will propose a comprehensive ethical framework for developing and evaluating virtual twin technology in liver disease. It will define ethical guidelines, including assessment frameworks, for project participants to verify compliance with ethical standards related to patient consent, privacy, data security and responsible use of AI. *Task 1.5 - Ethical Requirements* will oversee compliance with the ethical guidelines defined here, thereby ensuring patients' rights and promoting an ethical approach to VT-based patient management solutions by design and by default.

1. ETHICS GUIDANCE

The ARTEMIs activities will be conducted in full compliance with ethical principles and according to the European and national legal frameworks of the involved countries.

The ethical dimensions of the Project relate to its proposed **methodology** to conduct the ARTEMIs cohort study and to undertake R&D activities using patient data, including AI developments and the evaluation of VT models using patient data. The ethical dimensions of the Project **objectives** relate to the design of the CDSS system, its proposed features and operational workflows.

The ethics guidance will ensure that the Consortium complies with the specific rules related to Ethical aspects specified in the **Grant Agreement** (see Annex 1). These rules relate to factors such as the management of sensitive information, research integrity in the execution of project activities, and the protection of fundamental human rights.

1.1. General ethics considerations

The ARTEMIs project involves the secondary use of retrospective health data gathered as part of routine clinical practice and past observational studies, clinical trials, or research projects. The ARTEMIs project team is aware that using health data raises significant ethical issues concerning patients' privacy rights.

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

The Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons rules the processing of personal data and the free movement of such data (General Data Protection Regulation, GDPR). The project activities involving personal data are governed by the national implementations of the GDPR in the country of the participant entity.

Other International and European regulations/guidance to be observed include the following:

- The Charter of Fundamental Rights of the European Union of 12 December 2000.
- The Convention on Human Rights and Biomedicine of the Council of Europe (Oviedo, 04. IV, 1997).
- Additional Protocol to the Convention on Human Rights and Biomedicine, concerning Biomedical Research (Strasbourg, 25.I.2005).
- The revised World Medical Association Helsinki Declaration (Fortaleza, 2013).
- Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in conducting clinical trials on medicinal products for human use".
- ICH-GCP Guidelines, Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95), Sept. 1997.
- WHO Operating Guidelines for Ethics Committees that Review Biomedical Research, Geneva, 2000.

Project Data Protection Officer (DPO)

The jurist and DPO at MEDEX- BCP is the ARTEMIs Project DPO.

Project Board of DPOs, jurists and legal contacts

The DPO constitutes this Board at each data holder and data user entity of the ARTEMIs cohort. The full DPO Board members list, as well as other jurists or legal contacts who will be involved in the project, is available to all the consortium for easy consultation.

1.2. ARTEMIs Cohort

General description

The ARTEMIs retrospective cohort responds to the definition of a “retrospective collection and analysis of health data obtained from individual patients or healthy persons to address scientific questions related to the understanding, prevention, diagnosis, monitoring or treatment of a disease, mental illness, or physical condition” as defined in the work programme of this call. As such, the definition of a clinical study as defined by Regulation 536/2014 (on medicinal products) is not applicable in the framework of our study.

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

The ARTEMIs cohort is a multicentric cohort involving 12 university hospitals and medical research centres in 7 countries, corresponding to the following four clinical cases, wherein theory-based mechanistic and data-driven AI models will be developed and validated:

- **Clinical use case 1:** Liver disease staging in MASLD patients - Prediction model of disease fibrosis changes (progression and regression), with the ability to distinguish between fast and non-fast fibrosis progression among MASLD patients.
- **Clinical use case 2:** MASLD and progression of cardiovascular diseases.
- **Clinical use case 3:** Prediction of clinical outcomes in patients with cirrhosis and portal hypertension who receive TIPS placement or liver transplantation.
- **Clinical use case 4:** Management of Hepatocarcinoma (HCC) patients.

This cohort will serve the following main objectives:

- To develop and validate machine-learning or mechanistic models that can predict the evolution of liver diseases, particularly MASLD, at various stages (use cases 1 and 2) and the outcome of some intervention or therapies (use cases 3 and 4).
- To identify new correlations or validate suspected correlations between specific observations, interventions, and outcomes, particularly cardiovascular complications.

Consequently, the ARTEMIs cohort is divided into the following subgroups:

- **Clinical Use Cases 1 and 2:** Adult patients diagnosed with MASLD/MASH (based on clinical, biochemical and radiological features), with a minimum follow-up of 5 years.
  Clinical Use Case 2, a subgroup of patients with cardiovascular disease: Adult patients diagnosed with cardiac fibrosis (based on histological and radiological features), with or without MASLD, with a minimum follow-up of 1 year. This subgroup is required to study fibrosis mechanisms in cardiac tissue.

- **Clinical Use Case 3:** Adult patients diagnosed with cirrhosis (based on clinical, laboratory, endoscopic, ultrasonographic features or histology) and portal hypertension who receive TIPS placement or liver transplantation.

- **Clinical Use Case 4:** Adult patients diagnosed with HCC (based on radiology, cytology, or histology features) of any aetiology, who underwent HCC treatment (surgical interventions, ablation, TACE, TARE, SIRT and immunotherapies), with a minimum follow-up of 6 months after treatment.

- **Control group:** Adult patients with neither liver conditions nor cardiovascular events, with a minimum follow-up of 2 years, are used as healthy controls for the ARTEMIs models.

The cohort coordinator is Vall d'Hebron Barcelona Hospital (VHIR) in Spain. A complete description of ARTEMIs retrospective cohort can be found in the Study Protocol (Annex 2).

Throughout this report, *participants in the Artemis cohort* are also called *Data Holders.*

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

Ethics considerations

Each data holder and data user entity participating in the ARTEMIS Project has protocols in place to ensure compliance with their national implementations of the European General Data Protection Regulation (GDPR) for collecting, storing, protecting, retaining and destroying personal data. Table 2 indicates the national implementations of the GDPR in the countries where this Project is executed and the national data protection authority of reference for each project partner.

Table 1 – National GDPR implementations at the ARTEMIs participants´ country

Country	Participant (role)	National law	National data protection authority
France	*ICAN (health data holder)* *APHP (health data holder)* *INRIA (health data user)*	*Loi n° 2018-493 du 20 juin 2018 relative à la protection des données personnelles* - Law No. 2018-493 of 20 June 2018 on the protection of personal data.	Commission Nationale de l'Informatique et des Libertés (CNIL)
Spain	*MAT (no access to health data)* *VHIR (health data holder)* *HULAFE (health data holder)*	*Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales* - Organic Law 3/2018 of 5 December on the Protection of Personal Data and Guarantee of Digital Rights.	Agencia Española de Protección de Datos (AEPD)
Germany	*JUH (health data holder)* *UKHD (health data holder)* *CHARITÉ (health data holder)* *DKFZ (health data user)* *ALU-FR (health data user)* *ULEI (health data user)*	*Bundesdatenschutzgesetz (BDSG)* - Federal Data Protection Act.	Bundesbeauftragter für den Datenschutz und die Informationsfreiheit (BfDI)
Belgium	*CUSL (health data holder)* ELPA (no access to health data)	*Loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel* - Law of 30 July 2018 on the protection of natural persons with regard to the processing of personal data)	Autorité de protection des données (APD) / Gegevensbeschermingsautoriteit (GBA)
Austria	*MUV (health data holder)*	*Datenschutzgesetz (DSG)* - Data Protection Act	Datenschutzbehörde (DSB)
Italy	*ULS (health data holder)*	*Codice in materia di protezione dei dati personali* - Legislative Decree No. 196/2003 as amended by Legislative Decree No. 101/2018	Garante per la protezione dei dati personali (GPDP)
UK	ICL *(health data holder)*	Data Protection Act 2018 (DPA 2018)	Information Commissioner's Office (ICO)
CZECH R.	BETT (no access to health data)	*Zákon 110/2019 o zpracování osobních údajů* - Act No. 110/2019 on the Processing of Personal Data	Úřad pro ochranu osobních údajů (ÚOOÚ)
ISRAEL	SHEBA *(health data holder)*	own data protection law *Privacy Protection Law, 5741-1981*	Privacy Protection Authority (PPA)

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

Concerning the ARTEMIs cohort, the Project will use a hybrid federated/central infrastructure, with a data exploration portal for authorised users (work led by MEDEX-BCP) to perform data analytics and federated learning capabilities based on the Flower framework (work led by BU). For mechanistic models and AI tools that require it, a subset of cases will be transferred to a centralised cloud-based storage managed by MEDEX-BCP.

The data holder will anonymise patient data before sharing it with other project partners, either under the federated environment or in the central cloud-based storage. Thus, no directly identifiable data will be processed as part of the Project to comply with regulations and data confidentiality. Anonymisation of data is defined as the operation of irreversibly removing the personal character from a piece of data, considering all the means that are reasonably likely to be used to identify the natural person directly or indirectly (recital 26 of the GDPR).

After signing the required licensing agreement, MEDEX-BCP will provide the data holders with tools for anonymising imaging and clinical data. The anonymisation process will run as a double pseudonymisation process, where one corresponding key is removed once a dataset is created to ensure full anonymisation (more details can be found in the deliverable D2.1—Data Management Plan). MEDEX has successfully implemented this process in other Horizon R&D projects, such as ChAImeleon.

Each legal entity will designate authorised users within their site to use these tools. These tools will gather data from different sources (PACS, RIS, EHR Laboratories, etc.), anonymise it, and provide only meaningful data for the research in compliance with the principle of data minimization (Art.5 GDPR).

Data Sharing Agreement (DSA)

A DSA will regulate the collaboration related to data sharing and access between the project participants during its execution period. The DSA will define the roles and responsibilities of the data controllers / joint controllers and data processors. A model DSA is being finalised and submitted to the Consortium's institutional DPOs for review. The signed DSA is part of the dossier needed for the institutional Ethics Committee´s approval of some ARTEMIs cohort participants.

1.3. ARTEMIs CDSS

General description

The ARTEMIs project aims to deliver a functional prototype of a Clinical Decision Support System (CDSS) for a more personalised clinical management of MASLD patients in the context of the four use cases described in the previous section. The CDSS will be evaluated by a multi-country and multidisciplinary clinical panel (e.g., hepatologists, radiologists, pathologists, cardiologists, oncologists, and surgeons) as part of the Project.

ARTEMIs CDSS will be designed to provide an instant overview of the patient's multimodal data and exploit integrated virtual twin models to assist in predicting the evolution of the disease and cardiovascular outcomes, the response to a specific treatment, intervention or simply better life habits. This will be achieved through the integration and orchestration of multiscale (from molecular to entire organ), multilevel (signal transduction, metabolism, tissue mechanics, blood

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

flow and transport) and multiorgan (liver and cardiovascular system) virtual twin models. All in all, the system will contribute to the personalised management of these patients.

The modular CDSS prototype, illustrated in Figure 1, will consist of a visualisation module to display relevant information and a decision support module, where the user can call use-case-specific 'virtual twin' models, depending on the stage of the disease and the questions to be addressed.

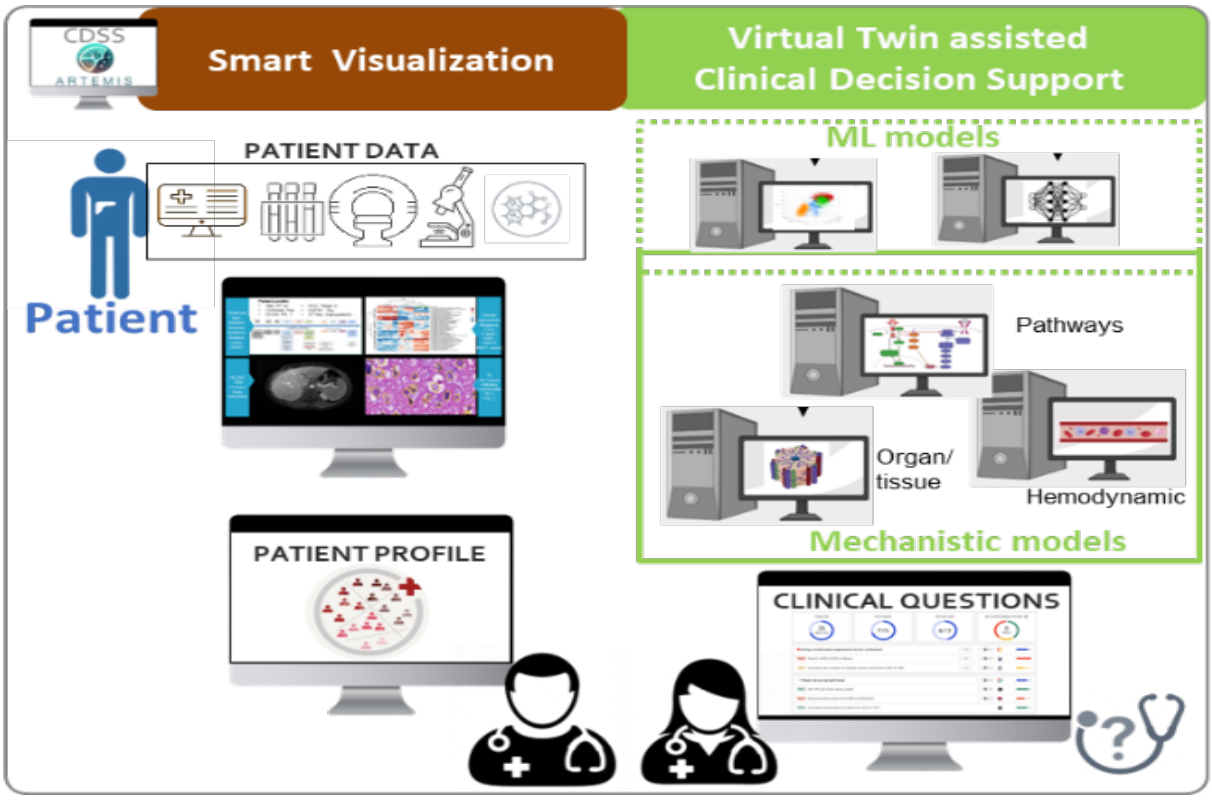

Figure 1: ARTEMIs modular CDSS

Ethics considerations

The CDSS aims to deliver clinically meaningful outcomes for the specific use cases proposed in MASLD patient management, providing robust results with mechanisms explained and confidences, biases, and hypotheses transparently disclosed. The proposed CDSS uses both mechanistic and AI models. The AI developments will fully align with the requirements for trustworthy AI established by the AI Act proposal to ensure appropriate performance and usability of models for their intended use in guiding clinical decisions.

To this aim, the project developments will use methodological frameworks that facilitate the development of trustworthy AI models (i.e. ALTAI tool). The ARTEMIs methodological approach will address challenges such as systematic discrimination by gender, age or other factors. The data to be used as training and testing datasets for AI developments will undergo thorough data analysis by the AI developer teams, using descriptive statistics, data visualisation, and data quality assessment before their use in the models. The fairness of the data will directly impact the fairness of the models, which is evaluated by the trustworthy AI frameworks (by the AI developers) and then assessed by the clinical experts in their validation works. Thus, several cycles of evaluation and refinement will be conducted, ensuring that fairness is evaluated at multiple levels (starting

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

with the data, continuing through model development, and concluding with clinical validation), enhancing the robustness and reliability of the results.

In their future path to market, the ARTEMIs CDSS will undergo a regulatory approval procedure as high-risk AI systems for use in healthcare (AI Act proposal) under the new regulation for Medical Devices. These requirements will be contemplated from the CDSS's design phase to ensure compliance and generate helpful documentation from the early stages.

As a system devoted to assisting and guiding clinical decisions, the design and implantation of the ARTEMIs CDSS functionalities will be aligned with ethical principles such as the respect for the human autonomy of the system users, the prevention of harm to the patients, and the delivery of explicability about the system outcomes. The system will respect human autonomy by design, as it is intended to assist decisions by providing additional meaningful clinical data. Nevertheless, the clinical professionals will oversee the decision-making process. The system will be transparent regarding models´ performance and confidence levels, using metrics and data visualisation methods that are meaningful to a clinical user. The CDSS dashboard will provide interfaces informing the data type, hypothesis, and other factors contributing to the prediction/simulation. For instance, the system will display the input data for the patient under study, which the model(s) has used to deliver the prediction/simulation. The clinical validation process during the project will focus on ensuring that the system's performance aligns with established ethical principles. It is important to emphasise that these validation activities will not influence or alter the real-time management of patients. Instead, the clinical validation will involve retrospective data (historical patient records) where clinical teams will simulate diagnostic or therapeutic decision-making processes with and without the support of the CDSS. The results from these processes will then be compared against each other and the original decisions made by the clinical teams at the time the patient was treated. This allows evaluation of the CDSS prototype´s performance of standard care methods without affecting patient care.

2. ETHICS IMPACT ASSESSMENT (EIA)

The ARTEMIs EIA is designed to **evaluate** the ethical implications of the Project in terms of its impact on the stakeholders in aspects such as fairness, justice, and societal rights. The following ethical considerations will be analysed:

- **Impact on privacy:** the actions to supervise and guide these ethical considerations are in place, ensuring data privacy in all project activities, as described in section 2.1.
- **Impact on personal autonomy, fairness, and transparency:** The CDSS's performance can impact these aspects due to the use of AI to support healthcare professionals in their decision-making processes. Assessing the trustworthiness of the developed AI system under a multifaceted vision is crucial. The planned actions are described in section 2.2.
- **Impact on patient rights, potential harm, and environmental impact:** The proposed CDSS's ethical implications are part of the holistic criteria guiding the recently started design process. The compliance revision will follow the design, prototyping, and evaluation advancements.

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

2.1. Compliance with the GDPR

The Consortium will put in place the required measures to implement the General Data Protection Regulation (GDPR), including:

Data pseudonymisation / Data anonymisation

Description of the proposed approach

To comply with regulations and data confidentiality, no directly identifiable data will be processed as part of the Project, and data will be **anonymised**, which should be distinguished from pseudonymisation:

- **Pseudonymised data** is legally considered personal data and, therefore, falls within the scope of the GDPR. Processing that separates identifying data from non-identifying data and uses a register of pseudonyms to link these two data sets is known as pseudonymisation.
- **Anonymised data** is not considered personal data because it does not allow the data subject to be identified. Anonymisation is the irreversible removal of the individual character from a piece of data.

To determine whether a natural person is identifiable, an account must be taken of all the means reasonably likely to be used to identify the natural person directly or indirectly (recital 26 of the GDPR).

The Clinical Centre side hosts patient data that may be identified. Each legal entity will designate authorised users within their site and will use the different provided tools by MEDEX to gather the data from the various data sources (PACS, RIS, EHR Laboratories, etc.), de-identify in a standardised way and provide only meaningful data for the research in compliance with the principle of data minimisation (Art.5 GDPR). The anonymisation process will run as a double pseudonymisation process, removing one corresponding key once a dataset is created to ensure complete anonymisation (Figure 2). This will be applied to both clinical and imaging data. MEDEX has successfully implemented such a process in the Horizon R&D ChAImeleon project. The advantage of this process is to ensure correspondence at the data preparation stage on-premise so that the clinical team can go back to identifying data source to complete missing information but to ensure that no.

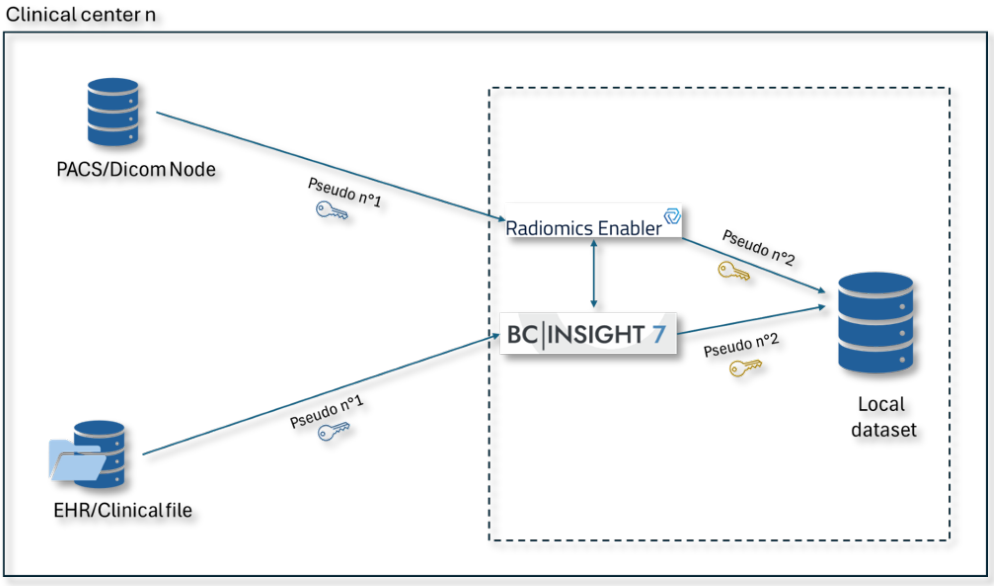

Figure 2: Two-step anonymisation process at one clinical site "n"

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

By default, a first pseudonymisation step is performed to replace the patient's name by a list of 10 random characters as pseudonym to each patient. The same pseudonym will be applied to the data of the corresponding patient (clinical and imaging). If the site provides already pseudonymised data, its team may choose to keep the original pseudonym for each patient or replace it.

For the second pseudonymisation step, additional encryption will be applied, with the hash of pseudonyms using sha256. Once double-pseudonymized datasets have been successfully exported to the local dataset, the key for the second pseudonymisation will be removed, ensuring complete anonymisation.

Description of the tools/processes to be used

MEDEX offers the deployment of tools for data collection, anonymisation, and harmonisation on-premise at the Data Holder site, free of charge, to facilitate data sharing. Installing these tools requires the signature of a license agreement for the duration of the project.

Clinical data of each patient will be collected in electronic clinical record forms (eCRF) under a module of **BC|Insight 7 tool**. This tool will allow data to be structured and harmonised across all clinical sites and use cases. Clinical data collection will be performed by manually editing eCRFs through a user interface for sites without existing datasets. For sites with already structured datasets, such datasets may be ingested into the BC|Insight 7 tool, providing they match, to some extent, the structure of the target eCRF (English language, similar naming of the clinical variables, etc.). Existing datasets may be provided in tabular (.csv, .xls) or wire (.json) format.

I**maging data** will be collected using the **Radiomics Enabler** tool, provided they are in DICOM format. This tool will also be deployed on-site on a second virtual machine (VM1). It will allow the extraction of radiological and/or digital pathology (slide macroscopy) imaging series directly from the hospital PACS. Alternatively, the tool may be connected to a specific DICOM node.

Data Protection Impact Assessment (DPIA)

The DPIA aims to identify threats and risks that may affect the right to data protection and the fundamental rights of data subjects. To assess the level of risk to these rights and freedoms, the DPIA needs to contemplate the complete data lifecycle in the project.

The DPIA is a mandatory process under Article 35 of the GDPR; thus, it is a must for entities using pseudonymised data. Although the DPIA is not compulsory for entities using anonymised data, it is a good practice to evidence the robustness of the data protection in this Project.

Data holders and users may have their own preferred DPIA template. To mitigate this situation, we have been working on a European common DPIA template that could apply to each country of the consortium.

This template is being finalised to fully match the project's data workflow and include the ad-hoc risk assessment. The respective roles on the project—data processor and data controller—are pending alignment with the data-sharing agreement. Each data holder and data user will revise the DPIA model to ensure it covers their processes and correctly reflects their role in the overall

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

workflow. The controller is the main actor in a DPIA. The processor must support the controller when conducting a DPIA.

The deliverable D1.4 Complete verification of obtaining approvals for clinical studies will include the results of the initial DPIA, corresponding to early identification of risks in managing patient data and their probability of occurrence for the data workflows designed in this Project.

To undertake this assessment, the critical steps of the data workflow must be defined, and the tools used and the responsible partners at each step must be identified. The DPIA will be an evolving document that will be revised and updated as the workflows are implemented and refined.

Secure IT infrastructure, traceability of data access and use

Hybrid architecture for data sharing: A hybrid architecture will be used for the project, allowing federated processing of data stored locally at the Data Holder's site, as well as the transfer of data to a central cloud platform provided by MEDEX, if required for specific data subsets or if this is the preferred option for some Data Holders. The description of the data workflow is provided in Annex 1. This workflow may be subject to updates during the project, which would be documented in the Data Management Plan (DMP).

Use of stringent security protocols: MEDEX tools for data pre-processing and storage support a wide range of data backup, recovery and retention options. Data collection tools are deployed according to the Data Holder site's security needs and recommendations using native cloud services. The tools encrypt all data in transit and support data encryption at rest. The method of encryption is based on the underlying infrastructure environment.

2.2 AI Impact Assessment

In July 2020, the European Commission's High-Level Expert Group on AI (HLEG-AI) published the Assessment List for Trustworthy Artificial Intelligence (ALTAI) tool. This tool enables organisations to perform self-assessments of the fit of their AI systems and surrounding governance to the "7 Principles for Trustworthy AI." The ALTAI assessment framework is recommended in major European initiatives to facilitate the use of health data in AI-based research and innovation, such as the Europe Cancer Image infrastructure.

ALTAI methodology helps AI developers from academia and the industry assess the trustworthiness of the AI models that they have developed. A web portal offers free access to an online tool.1 This tool facilitates the implementation of the self-assessment checklist, which contemplates seven domains:

- human agency and oversight
- technical robustness and safety
- privacy and data governance
- transparency
- diversity, non-discrimination and fairness
- environmental and societal well-being and
- accountability

1 https://altai.insight-centre.org/Home/HowToComplete)

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

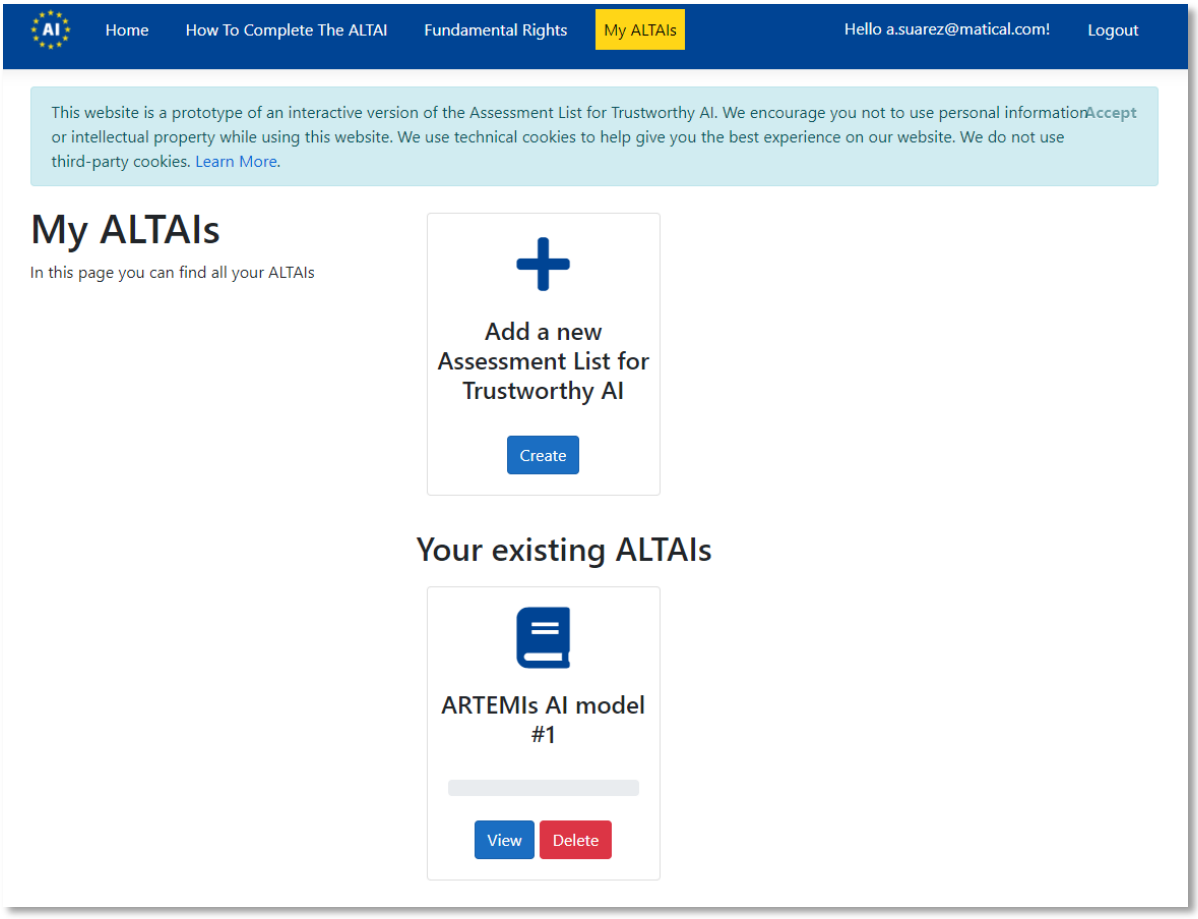

Figure 3: ALTAI self-assessment web portal

The ARTEMIs developers of AI-based models will be requested to complete an ALTAI assessment before their models are integrated in the CDSS prototype. They will use the online tool (Figure 3). The ALTAI self-assessment results and recommendations will be informed to MAT as leader of *Task 1.5 - Ethical Requirements*. The assessment outputs will be jointly analysed with the developer, paying particular attention to assessments that output low scores, indicating an inadequate performance in some domains. A plan for model refinements will be elaborated.

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

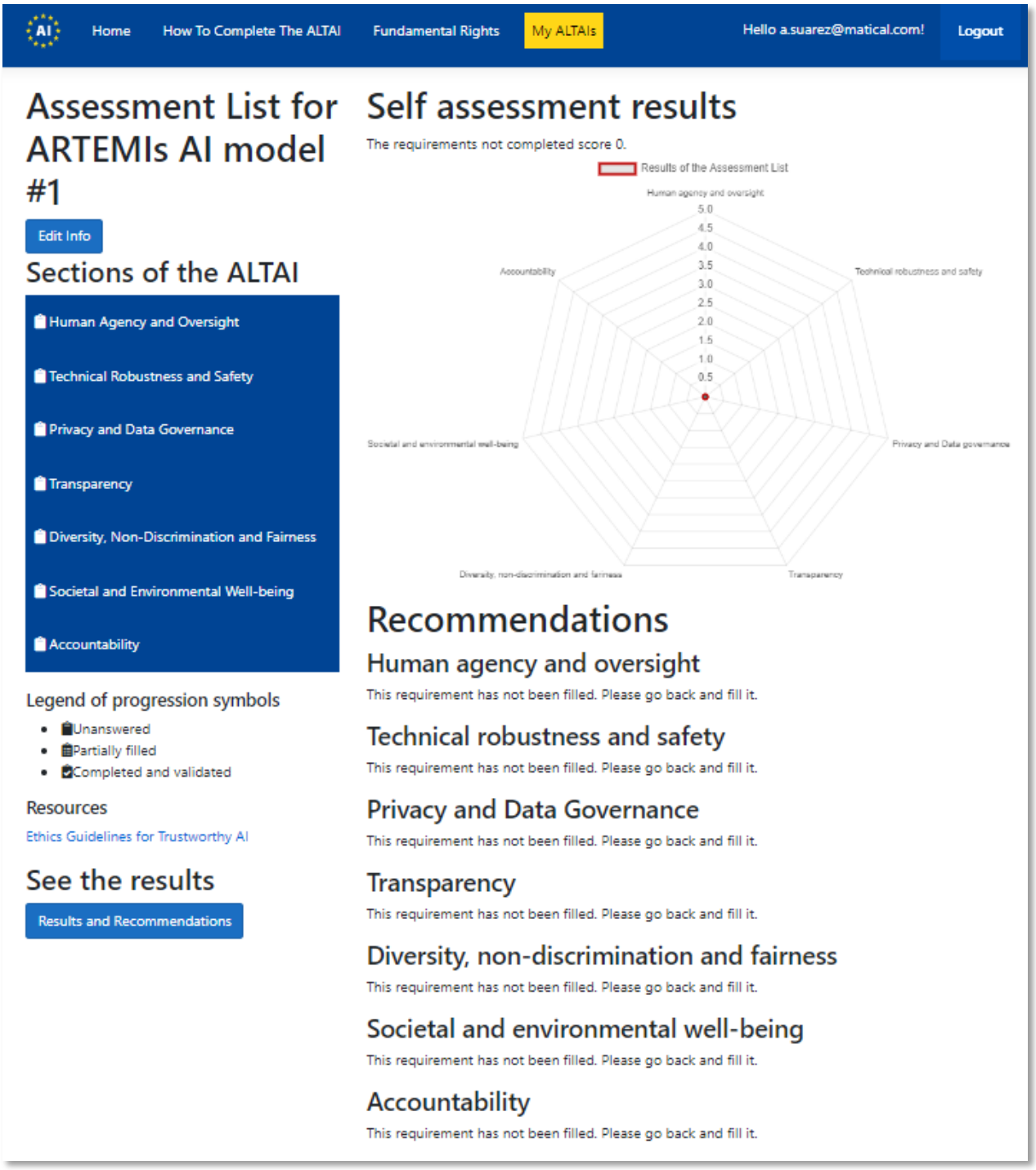

Figure 4: Radar of scores per domain and list of recommendations generated by the ALTAI self-assessment

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

3. INSTITUTIONAL ETHICS COMMITTEE APPROVALS

A study protocol has been elaborated for the retrospective phase of the ARTEMIs cohort. The study protocol comprehensively describes the primary and secondary objectives, the study population, inclusion and exclusion criteria, and procedures for conducting the study for each of the proposed Use Cases (UC). The data deidentification, storage, and sharing methods will be elaborated further in the Data Management Plan (DMP). The first version of the DMP (Deliverable D2.1) was prepared by MEDEX and submitted in month 6. The study protocol was elaborated as part of WP2, under the coordination of MEDEX and VHIR, in close collaboration with the UC clinical and technical leaders.

Additionally, for a subset of volunteer patients, the prospective collection of biological samples and clinical data is also planned for future stages of the project. A revised study protocol will be prepared for the prospective phase, where aspects such as the procedures for collecting biological samples (when applicable) and details on the duration of storage will be defined. The Patient Informed Consent form (ICF) and patient information sheet will be used for prospective data and sample collection. Before requesting the patient to sign the ICF, the Medical Doctor will verify that the patient is well informed, as explained in the Patient Information Leaflet (in language and terms intelligible to the participants). The ICF will be written according to the EU Regulation and good clinical practice in clinical trials.

3.1 Requirements for each participant in the ARTEMIs Cohort

The data holders in the ARTEMIs cohort have informed their institutional DPOs and identified the legal and ethical compliance requirements specific to their institutional Ethics Committee (iEC) approval request process. These requirements vary under their respective national and regional data protection regulations and ethical guidelines.

Some universities, such as UKHD or JUH, have published the list of documents required for ethical submission at their own facility (only available in German). However, this list always has to be adapted to the context of the research project.
Each site is, in principle, responsible for submitting to the iEC, and all have experience in this process from previous research projects. However, whenever the sites requested assistance to collect the appropriate documents or discuss the submission preparation, MEDEX has offered to help them and ensure they get all the information they need.

Table 3 below summarises the information we have for each participant in the Cohort. As all sites submit to the iEC themselves, we do not always understand what is required for each submission. Some information comes from conversations with partners.

Table 2 – List of requirements per participant in the ARTEMIS cohort

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

ORGANIZATION	Country	requirements for ethics
HULAFE	Spain	Authorization request form Project record (project summary, study description, project budget) Data protection and law compliance statement
VHIR	Spain	Study protocol per use case Commitment letter from the PI and collaborators Submission fee invoice form iEC Evaluation request Request for exemption from informed consent Ethical approval from another Spanish site (HULAFE)
ICAN	France	*NA* *
AP-HP (Mondor H.)	France	*NA* *
AP-HP (Salpêtrière)	France	*NA* *
AP-HP (Paul Brousse)	France	*NA* *
CUSL	Belgium	Study protocol GDPR questionnaire (containing : information on patient ID management; information on software deployment on site; data storage and protection plan) List of clinical variables Declaration of costs signed by PI Data sharing agreement (draft)
JUH	Germany	General project concept List of all data to share Technical and organizational measures (TOM) Data protection concept Deletion concept (part of Data Management Plan) Detailed description of the data flow (part of Data Management Plan) Concept of what data is required and how it is to be transferred (part of Data Management Plan) Requirements for which IT is needed and how it should be used Waiver request for informed consent
UKHD	Germany	Study protocol List of Clinical Data Data Sharing Concepts
CHARITE	Germany	Study protocol Information letter Consent form List of clinical data to share Ethical approval from another site
ULEI	Germany	Study protocol List of clinical data to share Data sharing agreement and material transfer agreement
ULS	Italy	[information pending]
MUV	Austria	Study protocol List of clinical variables Waiver for informed consent
ICL	UK	[information pending]

*Sites in France need nonetheless to go through a few regulatory steps: 1/ internal validation of the applicable methodology of reference (MR004); 2/Project declaration to the CNIL; 3/ information letter to patients

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

3.2 Current status of iEC´s approval for each participant in the ARTEMIs Cohort

Table 4 summarises the status of this request process for each participant in the Cohort as of the date of submission of this report. The PI at each participant is responsible for ensuring that only patient data is incorporated into the ARTEMIs cohort storage if ethical approval is awarded and kept on file. The PI at each participant is also responsible for maintaining a file of all final approval documents. The Cohort Coordinator (VHIR) monitors the progress of the ethics approval request at each participant to mitigate potential delays.

Table 3 – Ethical approval status for each participant in the ARTEMIS cohort

ORGANIZATION	Country	Ethics status
ICAN	France	No ethical committee submission required, pending internal validation
AP-HP (Mondor H.)	France	No ethical committee submission required, pending internal meeting
AP-HP (Salpêtrière)	France	No ethical committee submission required, pending internal meeting
AP-HP (Paul Brousse)	France	No ethical committee submission required, pending internal meeting
JUH	Germany	Pending internal meeting to review what is missing for ethical submission Need more insights internally of data to be collected
VHIR	Spain	Plan to submit in October 2024, pending data collection description, and study protocol split per use cases
UKHD	Germany	Clinical team working on compiling the files for submission, they will need the study protocol split per use cases
CUSL	Belgium	Ready to submit, once the Data Sharing Agreement model is available
HULAFE	Spain	Ethical approval obtained for all Use Cases; Pending DPIA (Data Protection Impact Assessment)
CHARITE	Germany	Pending internal meeting to review what is missing; would benefit from approval from another site (German or any other; have the one from HULAFE)
ULS	Italy	Clinical team working on compiling the files for submission. Will submit the study as an amendment to another research study that already received ethical approval.
MUV	Austria	Clinical team working on compiling the files for submission.
ICL	UK	Ask assistance for ethical submission, pending a new meeting with MEDEX
ULEI	Germany	Clinical team working on compiling the files for submission, and asked for a material transfer agreement (potential annex to the DSA).

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

4. CONCLUSION

This document provides an overview of how compliance with ethical principles will be approached in executing the ARTEMIs Project. The ethics guidance has been designed to address the specific conditions of the ARTEMIs' main components: the cohort and the CDSS. A detailed plan has been defined for the conductance of ethics impact assessment that effectively addresses identifying and mitigating the main existing risks. The ethics guidance aims to ensure legal and ethical compliance, not only in the activities undertaken during the project but also in the functionality of the results to be achieved.

5. FUTURE WORK

Future work corresponds to following the roadmap defined in this report, gathering the required information and knowledge to elaborate on the required ethics impact assessments in the terms that have been described.

The main planned following works are:

- Have a DSA signed by all Consortium members.
- Have a Material Transfer Agreement (when applicable) signed with the involved participants.
- Have a finalised version of DPIA to guide the assessment of the Cohort participants.
- Monitor the progress in the ethics approval process and the DPIA conductance of the cohort participants.
- Supervise compliance with the ethics guidance in the project activities.
- Supervise compliance with ethical considerations in the design of the CDSS.
- Each developer is to complete the ALTAI self-assessment for their AI model(s) before the end of the Project.

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

ANNEX 1

ARTEMIs GRANT AGREEMENT ANNEX 5

Specific Rules Related to Ethics (Article 14)

ANNEX 5

SPECIFIC RULES

CONFIDENTIALITY AND SECURITY (— ARTICLE 13)

Sensitive information with security recommendation

Sensitive information with a security recommendation must comply with the additional requirements imposed by the granting authority.

Before starting the action tasks concerned, the beneficiaries must have obtained all approvals or other mandatory documents needed for implementing the task. The documents must be kept on file and be submitted upon request by the coordinator to the granting authority. If they are not in English, they must be submitted together with an English summary.

For requirements restricting disclosure or dissemination, the information must be handled in accordance with the recommendation and may be disclosed or disseminated only after written approval from the granting authority.

EU classified information

If EU classified information is used or generated by the action, it must be treated in accordance with the security classification guide (SCG) and security aspect letter (SAL) set out in Annex 1 and Decision 2015/444[1] and its implementing rules — until it is declassified.

Deliverables which contain EU classified information must be submitted according to special procedures agreed with the granting authority.

Action tasks involving EU classified information may be subcontracted only with prior explicit written approval from the granting authority and only to entities established in an EU Member State or in a non-EU country with a security of information agreement with the EU (or an administrative arrangement with the Commission).

EU classified information may not be disclosed to any third party (including participants involved in the action implementation) without prior explicit written approval from the granting authority.

ETHICS (— ARTICLE 14)

Ethics and research integrity

The beneficiaries must carry out the action in compliance with:

- ethical principles (including the highest standards of research integrity)

---

[1] Commission Decision 2015/444/EC, Euratom of 13 March 2015 on the security rules for protecting EU classified information (OJ L 72, 17.3.2015, p. 53).

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

and

- applicable EU, international and national law, including the EU Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms and its Supplementary Protocols.

No funding can be granted, within or outside the EU, for activities that are prohibited in all Member States. No funding can be granted in a Member State for an activity which is forbidden in that Member State.

The beneficiaries must pay particular attention to the principle of proportionality, the right to privacy, the right to the protection of personal data, the right to the physical and mental integrity of persons, the right to non-discrimination, the need to ensure protection of the environment and high levels of human health protection.

The beneficiaries must ensure that the activities under the action have an exclusive focus on civil applications.

The beneficiaries must ensure that the activities under the action do not:

- aim at human cloning for reproductive purposes
- intend to modify the genetic heritage of human beings which could make such modifications heritable (with the exception of research relating to cancer treatment of the gonads, which may be financed)
- intend to create human embryos solely for the purpose of research or for the purpose of stem cell procurement, including by means of somatic cell nuclear transfer, or
- lead to the destruction of human embryos (for example, for obtaining stem cells).

Activities involving research on human embryos or human embryonic stem cells may be carried out only if:

- they are set out in Annex 1 or
- the coordinator has obtained explicit approval (in writing) from the granting authority.

In addition, the beneficiaries must respect the fundamental principle of research integrity — as set out in the European Code of Conduct for Research Integrity[2].

This implies compliance with the following principles:

- reliability in ensuring the quality of research reflected in the design, the methodology, the analysis and the use of resources
- honesty in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair and unbiased way

[2] European Code of Conduct for Research Integrity of ALLEA (All European Academies).

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

- respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment
- accountability for the research from idea to publication, for its management and organisation, for training, supervision and mentoring, and for its wider impacts

and means that beneficiaries must ensure that persons carrying out research tasks follow the good research practices including ensuring, where possible, openness, reproducibility and traceability and refrain from the research integrity violations described in the Code.

Activities raising ethical issues must comply with the additional requirements formulated by the ethics panels (including after checks, reviews or audits; see Article 25).

Before starting an action task raising ethical issues, the beneficiaries must have obtained all approvals or other mandatory documents needed for implementing the task, notably from any (national or local) ethics committee or other bodies such as data protection authorities.

The documents must be kept on file and be submitted upon request by the coordinator to the granting authority. If they are not in English, they must be submitted together with an English summary, which shows that the documents cover the action tasks in question and includes the conclusions of the committee or authority concerned (if any).

VALUES (— ARTICLE 14)

Gender mainstreaming

The beneficiaries must take all measures to promote equal opportunities between men and women in the implementation of the action and, where applicable, in line with the gender equality plan. They must aim, to the extent possible, for a gender balance at all levels of personnel assigned to the action, including at supervisory and managerial level.

INTELLECTUAL PROPERTY RIGHTS (IPR) — BACKGROUND AND RESULTS — ACCESS RIGHTS AND RIGHTS OF USE (— ARTICLE 16)

Definitions

Access rights — Rights to use results or background.

Dissemination — The public disclosure of the results by appropriate means, other than resulting from protecting or exploiting the results, including by scientific publications in any medium.

Exploit(ation) — The use of results in further research and innovation activities other than those covered by the action concerned, including among other things, commercial exploitation such as developing, creating, manufacturing and marketing a product or process, creating and providing a service, or in standardisation activities.

Fair and reasonable conditions — Appropriate conditions, including possible financial terms or royalty-free conditions, taking into account the specific circumstances of the request for access, for example the actual or potential value of the results or background to which access is requested and/or the scope, duration or other characteristics of the exploitation envisaged.

3

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

ANNEX 2

ARTEMIs COHORT – Study Protocol (V 30/09/2024)

This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

ARTEMIs : The Power of Virtual Twins to Fight MASLD

ARTEMIs RETROSPECTIVE COHORT PRE-PROTOCOL

Table of Contents

GLOSSARY

CT : Computed Tomography

CVD : Cardio-Vascular Disease

HCC : Hepato Cellular Carcinoma

MASH : Metabolic dysfunction-associated steatohepatitis

MASLD : Metabolic dysfunction-Associated Steatotic Liver Disease

MRI : Magnetic Resonance Imaging

PET : Positron Emission Tomography

SLD : Steatotic Liver Disease

TACE : Trans-Arterial ChemoEmbolisation

TARE : Trans-Arterial RadioEmbolisation

TIPS : Transjugular Intrahepatic Portosystemic Shunt

US : Ultrasound

USE : Ultrasound elastography

VCTE : Vibration-Controlled Transient Elastography

1. DESCRIPTION OF THE CLINICAL STUDY

ARTEMIs retrospective cohort responds to the definition of a "**retrospective collection and analysis of health data obtained from individual patients or healthy persons in order to address scientific questions related to the understanding, prevention, diagnosis, monitoring or treatment of a disease, mental illness, or physical condition**" as defined in the work programme of this call. In such, the definition of a clinical study as defined by Regulation 536/2014 (on medicinal products) is not applicable in the framework of our study.

The cohort will serve the following main objectives:

- To develop and validate machine-learning or mechanistic models that can predict the evolution of liver diseases, in particular MASLD at various stages (use cases 1 and 2), and the outcome of some intervention or therapies (use cases 3 and 4).
- To identify new correlations or validate suspected correlations between specific observations, interventions and outcomes, in particular cardiovascular complications.

COHORT COORDINATOR

Raul Herance (raul.herance@vhir.org), Vall d'Hebron Barcelona Hospital, Spain

TITLE, ACRONYM, UNIQUE IDENTIFIER

Accele**R**ating the **T**ranslation of virtual twins towards a p**E**rsonalised **M**anagement of fatty l**I**ver patients (ARTEMIs) Retrospective Cohort

STUDY RATIONALE

Metabolic dysfunction-associated steatotic liver disease (MASLD) is presently the most common chronic liver disease worldwide, accounting for a global prevalence of 25.24% (2). Its natural history remains unclear, given the multiple pathways through which disease progression takes place (3), as well as to the shortage of population-based studies addressing its long-term prognosis (4). As an attempt to alleviate the paucity of good quality data on MASLD's natural history (5) and to improve patient's care, the ARTEMIS project envisages to constitute a longitudinal cohort comprising patients at various stages of liver diseases, with emphasis on MASLD (use cases 1 and 2).

Given the remarkable heterogeneity underlying MASLD mechanisms, the deployment of computational models has increased in popularity among the scientific community, as an effective means to unravel this intricate subject (6). In particular, the understanding of the human liver metabolism plays a key role towards a deeper understanding of the main drivers that rule disease progression. In such, mechanistic models play a major role in the representation of the complexity that is inherent to the liver and the gastroenterology system. In a complementary way, machine learning models are expected to respond to more precise questions related to different stages of the disease and related comorbidities, therefore allowing the prediction of diagnosis and prognosis, as well as risk stratification, based upon parameters that are specific to each subpopulation.

In this light, the ARTEMIS cohort will be used to test new hypotheses, as well as to train, validate and evaluate the performance of computational models - including machine-learning models, mechanistic models and associations thereof - aimed to improve the management of MASLD patients. The ARTEMIs cohort will incorporate retrospective multisource data for MASLD patients along the spectrum of the disease, thus including MASH, cirrhosis and HCC patients. The cohort will include patients from 12 centres in 7 countries. The cohort will also incorporate data related to the most relevant comorbidities associated with these populations, most notably, cardiovascular events.

EXTENT AND EVALUATION OF CURRENT KNOWLEDGE DIRECTLY LINKED TO THE SCIENTIFIC QUESTION(S) TO BE ANSWERED BY THE CLINICAL STUDY

In addition to the complexities concerning its natural history, MASLD has been associated with an increased risk of developing cardiovascular disease (CVD) and cardiac events, including coronary artery disease, atherosclerosis, heart failure, and arrhythmia. The exact mechanism by which MASLD increases the risk of CVD is not fully understood, but it is thought to be related to the systemic inflammation and metabolic dysfunction associated with the condition.

Several studies have investigated the relationship between MASLD and cardiac events. A systematic review and meta-analysis published in 2016 (7), analysed 16 prospective and retrospective cohorts with 34,043 adult individuals (36.3% with MASLD) and approximately 2,600 CVD outcomes (>70% CVD deaths) over a median period of 6.9 years. They concluded that MASLD is associated with an increased risk of fatal and non-fatal CVD events, although the design of the observational studies did not allow to draw definitive causal inferences.

There is a consensus that MASLD patients should be closely monitored for cardiovascular risk factors and managed accordingly to reduce their risk of developing CVD. Nevertheless, given the high current prevalence of the disease and its expected growth, such monitoring may enormously stress the public healthcare systems.

Solutions that help to stratify those MASLD patients at higher risk of suffering cardiovascular events, are needed. The ARTEMIs cohort is aimed to assist the development of this type of solutions, based on advanced computational models.

OBJECTIVE(S) OF THE CLINICAL STUDY

The ARTEMIs project envisages to consolidate a holistic virtual model allowing, on the one hand, a better understanding of the underlying mechanisms involved in MASLD progression, as well as the prediction of cardiovascular events at different stages of the disease. In this light, 4 clinical cases will be considered, wherein theory-based mechanistic and data-driven AI models will be developed and validated, either individually or in association, depending on the clinical questions being raised.

The objective of ARTEMIs cohort is to assess the performance of mechanistic and AI-based models that will be deployed in the different clinical cases, based on their respective sensibility and specificity. The primary and secondary objectives regarding each clinical case study are presented in Table 1 herein below:

Table 1 : ARTEMIs clinical cases and related primary and secondary objectives

Clinical case	*Objectives*	*Outcomes*
Clinical Use Case 1: Liver disease staging in MASLD patients - Prediction model of disease changes (progression and regression), with ability to distinguish between fast and non-fast fibrosis progression among MASLD patients	**Primary objective** Evaluate the model's ability to distinguish progressors vs regressors or no changes in the disease severity spectrum.	**Primary Outcomes** Probability rate of disease changes (fibrosis and steatohepatitis - MASH) using validated non-invasive tests (NITs) imaging and liver histology, whenever available.

	Secondary Objectives:
	• Evaluate the model's ability to distinguish between fibrosis progressors and regressors.
	• Evaluate the model's ability to distinguish between fast and slow fibrosis progression.
	• Evaluate the model's ability to identify at-risk patients and time of developing clinically significant outcomes among the following:
	  ○ Progression to cirrhosis
	  ○ Liver decompensation
	  ○ HCC occurrence
	  ○ Liver transplantation
	  ○ Mortality (liver and non-liver related)
	Exploratory Objectives:
	• Better understanding of mechanism of actions of specific therapeutic interventions for MASLD associated comorbidities (i.e. obesity, T2DM or cardiovascular diseases) and their impact on the liver damage (fibrosis, steatohepatitis, MASH).
	• Evaluation of the correlations between lifestyle modifications, body composition (visceral obesity, sarcopenia) and liver damage (fibrosis stage/changes and steatohepatitis).
		Secondary outcome measures:
		• Probability rate of fibrosis progression and regression with distinction between fast and slow progressors using validated NITs, imaging, and liver histology when available.
		• Probability rates and time to progression to clinically significant outcomes:
		  ○ Time to progression to cirrhosis
		  ○ Time to liver decompensation
		  ○ Time to HCC occurrence
		  ○ Time to liver transplantation
		  ○ Time to mortality

<table>
<tr>
<td rowspan="2">Clinical Use Case 2:
MASLD and progression of cardiovascular diseases</td>
<td>Primary objective:
Assessment of a computational model to predict the incidence of major fibrosis-associated cardiovascular complications including syndromic, metabolic, and multicellular (heart) diseases. [Time Frame: Patients will be followed for an expected mean time of 5 years].</td>
<td>Primary outcome measure:
Any cardiovascular events (myocardial infarction, stroke, atrial fibrillation) or comorbidities.</td>
</tr>
<tr>
<td>Secondary objective:
Evaluate the impact of hepatic fibrosis and MASH (inflammation) on early cardiovascular risk and the degree of cardiac dysfunction.

Exploratory objectives:

• Evaluation of the impact of fibrosis in immunological system
• Identification of therapeutic targets
• Evaluation of response to treatment</td>
<td>Secondary outcomes measures:
• Framingham Cardiovascular risk score [Time frame: 36 months]
• Liver stiffness measured by vibration controlled transient elastography (VCTE)”
• Measurement of clinical, biochemical (fibrosis stage) assessment: fibroscan, liver histology, sheer wave elastography, validated biomarkers scores such as FIB4.
• Systemic Coronary Risk Estimation (SCORE) ranging from <1% very low risk to >15% very high risk of cardiovascular mortality.
• Coronary calcium score
• Coronary heart disease [Time Frame: 3 months] - Diagnosis of coronary heart disease, measurement of coronary flow.
• Evaluation of changes in left ventricular ejection fraction (EF) and hospitalization for heart failure.
• Determination of parameters of cardiac remodelling (e.g. myocardial wall thickness, myocardial radiodensity, epicardial adipose tissue).
• Prevalence of insulin resistance.
• Measurement of visceral fat in routine CT/MRI images.</td>
</tr>
</table>

<table>
<tr>
<td rowspan="2">Clinical Use Case 3-TIPS:
Computational models for the prediction of clinical outcomes in patients with cirrhosis and portal hypertension who receive TIPS placement</td>
<td>Primary objective:
Evaluate the performance (sensitivity, specificity, positive predictive value and negative predictive value, and likelihood ratios) of a predictive model (association of mechanistic and AI-based models) of cardiovascular complications for patients with cirrhosis undergoing a TIPS placement.</td>
<td>Primary Outcome measure
Incidence rate of heart failure[1] after insertion of TIPS. [Time Frame: from insertion of TIPS to 1 year post TIPS]

Patients will be classified in 3 grades: no heart failure, suspicion of heart failure (for those who have clinical suspicion but do not fulfill the defined criteria, or do not have the measurement of natriuretic peptides) and heart failure.</td>
</tr>
<tr>
<td>Secondary objectives:
• Evaluation of overall survival
• Prediction of further decompensation of cirrhosis (variceal bleeding, hepatic encephalopathy, ascites, jaundice and infections)
• Need for further paracentesis beyond 3 months after TIPS placement.

Exploratory Objectives:
• Better understanding of mechanism of actions of cardiac-related events after TIPS placement
• Evaluate the performance of the model in optimizing TIPS intervention</td>
<td>Secondary outcomes measures:
• Overall survival [Time Frame: all patients will be followed for 1 year after TIPS placement]- evaluation of overall survival from time of TIPS placement to death.
• Further decompensation according to the Baveno VII definition [Time Frame: from insertion of TIPS to 1 year post TIPS] – variceal bleeding, hepatic encephalopathy, ascites and its complications (hepatorenal syndrome, spontaneous bacterial peritonitis) and jaundice.
• Infections [Time Frame: from insertion of TIPS to 1 year post TIPS].
• Acute on chronic liver failure (according to the CLIF definition) decompensation [Time Frame: from insertion of TIPS to 1 year post TIPS].</td>
</tr>
</table>

Clinical Use Case 3-LV: Computational models for the prediction of clinical outcomes in patients with cirrhosis and portal hypertension who receive liver transplantation	**Primary objective:** Evaluate the performance (sensitivity, specificity, positive predictive value and negative predictive value, and likelihood ratios) of a predictive model (association of mechanistic and AI-based models) of cardiovascular complications for patients with cirrhosis undergoing liver transplantation.	**Primary outcome measures** Incidence rate of cardiac events after transplantation; including heart failure$_1$, myocardial infarction, symptomatic coronary heart disease, arrythmia. [Time Frame: from insertion of transplant to 5 years post-transplant].
	Secondary objective: Evaluation of overall survival - Evaluate the performance (sensitivity, specificity, positive predictive value and negative predictive value, and likelihood ratios) of a predictive model (association of mechanistic and AI-based models) of **early** cardiovascular complications (within 1 year) for patients with cirrhosis undergoing liver transplantation. - Evaluate the performance (sensitivity, specificity, positive predictive value and negative predictive value, and likelihood ratios) of a predictive model (association of mechanistic and AI-based models) of **late** cardiovascular complications (1-5 years) for patients with cirrhosis undergoing liver transplantation. **Exploratory Objectives:** - Better understanding of mechanism of actions of cardiac-related events after liver transplantation. - Evaluate the performance of the model in optimizing liver transplantation management.	**Secondary outcome measure** - Overall survival [Time Frame: all patients will be followed for 5 year after transplantation]- evaluation of overall survival from time of transplantation to death. - Incidence rate of early cardiac events after transplantation; including heart failure, myocardial infarction, symptomatic coronary heart disease, arrythmia. [Time Frame: from insertion of transplant to 1-year post-transplant] - Incidence rate of late cardiac events after transplantation; including heart failure, myocardial infarction, symptomatic coronary heart disease, arrythmia. [Time Frame: from insertion of transplant to 1-5 years post-transplant].

Clinical Use Case 4: Prediction of cardiovascular complications due to HCC treatments* **Note: includes surgical interventions, ablation, TACE, TARE, SIRT and immunotherapies*	**Primary objective:** Assessment of the feasibility of the predictive role of a computational model on the incidence of cardiovascular events related to therapeutic responses in HCC patients.	**Primary outcome measures** Incidence rate of cardiac-related events; including myocardial complications, heart failure and heart attack [Time Frame: 2 years]
	Secondary objective: Assessment of the onset of cardiac complications related to therapeutic strategies addressed to HCC patients, at different stages of the disease	**Secondary outcome measure** Delay of cardiac complications in relation to the diagnosis of HCC [Time Frame: 2 years].

[1]Definition of heart failure according to the European Society of Cardiology guidelines for the diagnosis and treatment of acute and chronic heart failure. In patients without history of heart failure, it was retained by the presence of typical symptoms of heart failure (e.g. breathlessness, ankle swelling and fatigue) accompanied by signs of heart failure (e.g. elevated jugular venous pressure, pulmonary crackles and peripheral edema) with elevated intracardiac pressures as assessed by increase of natriuretic peptides. Final diagnosis of severe cardiac decompensation was retained when leading to hospital admission for intravenous diuretic administration.

CHARACTERISTICS OF THE STUDY POPULATION

ARTEMIS cohort is divided in the following subgroups:

- **Clinical Use Case #1&2:** Adult patients diagnosed with MASLD/MASH (based on clinical, biochemical and radiological features), with a minimum follow-up of 5 years.
    - Clinical Use Case #2, subgroup of patients with cardiovascular disease: Adult patients diagnosed with cardiac fibrosis (based on histological and radiological features), with or without MASLD, with minimum follow-up of 1 year. This subgroup is required to study fibrosis mechanisms in cardiac tissue.
- **Clinical Use Case #3:** Adult patients diagnosed with cirrhosis (based on clinical, laboratory, endoscopic, ultrasonographic features or on histology) and portal hypertension who receive TIPS placement or liver transplantation.
- **Clinical Use Case #4:** Adult patients diagnosed with HCC (based on radiology, cytology, or histology features) of any aetiology, who underwent HCC treatment (surgical interventions, ablation, TACE, TARE, SIRT and immunotherapies), with a minimum follow-up of 6 months after treatment.
- **Control group:** Adult patients with neither liver conditions nor cardiovascular events with a minimum follow-up of 2 years, as healthy control for the ARTEMIs models.

In order to ensure early access to quality data, for the computational models' development and validation, we propose the use of longitudinal data on retrospective cases from existing cohorts available at our partners' sites, as well as retrospective cases gathered during standard clinical practice or clinics, at each participant clinical site.

A subset of liver tissue, plasma and serum samples, already available in the biobanks of existing cohorts, will undergo proteomics analysis performed by DKFZ and metabolomics and lipidomics analysis will be performed by VHIR.

Table 2 : ARTEMIS cohort and subpopulations (inclusion and exclusion criteria)

Clinical cases & subpopulations	***Inclusion and exclusion criteria***
Clinical Use Case 1 Liver disease staging in MASLD patients - Prediction model of fibrosis changes (progression and regression), with ability to distinguish between fast and non-fast fibrosis progression among MASLD patients.	**Inclusion criteria:** • Age ≥18 years • Diagnosis of MASLD confirmed by radiological imaging (any type: MR, CT, PET, VCTE, US, USE...) or histology (gold standard, following MASH SAF score) • With at least one follow-up of minimum 1 year after diagnosis of MASLD, with radiological imaging or histology
	Exclusion criteria: • Missing data on blood glucose, BMI and metabolic status. • Patients who have received systemic chemotherapy • Patients with hepatitis B (HBV) and hepatitis C (HCV), alcoholic liver disease (more than 5 years of drinking history, equivalent to alcohol volume ≥ 30g / D in male and ≥ 20g / D in female), drug-induced liver disease or autoimmune hepatitis. • Subjects having a significant risk of bleeding (platelet < 50x109 / L, prothrombin activity < 50%) • Presence of any other form of chronic liver, at the time of MASLD diagnosis.

Clinical Use Case 2: MASLD and progression of cardiovascular diseases	**Inclusion criteria:** • Age ≥18 years • MASLD patients regardless of disease stage of severity (from simple steatosis to cirrhosis) • Patients without known heart disease • Cardiovascular assessment available
	Exclusion criteria: • Association with another cause of liver disease • History of hepatitis B or C • Already known coronary artery disease • History of cardiovascular events
Clinical Use Case 3-TIPS: Patients with cirrhosis and portal hypertension who receive TIPS placement.	**Inclusion criteria:** • Age ≥18 years • TIPS indication (Baveno VII), except pre-emptive and salvage TIPS. – Recurrent variceal bleeding after failure of the usual pharmacological and endoscopic methods – Refractory or recurrent ascites or difficult to treat – Refractory Hydrothorax • Patients with diagnosis of liver cirrhosis (based on laboratory parameters, clinical, endoscopic, radiological or histological findings), of any aetiology.
	Exclusion criteria: • Portosinusoidal vascular disease • Complete portal vein thrombosis • Patients with surgical porto-caval shunts. • Patients with evidence of current locally advanced or metastatic malignancy • Patients with acute or chronic heart failure (New York Heart Association [NYHA]). • Patients with chronic obstructive pulmonary disease GOLD grade III/IV • Patients with chronic kidney disease requiring renal replacement therapy • Patients with a known infection with human immunodeficiency virus (HIV) or have clinical signs and symptoms consistent with current HIV infection • Patients with previous liver transplantation • Patients lost to follow-up and therefore have an incomplete 1-year follow-up

Clinical Use Case 3-LV: Patients with cirrhosis and portal hypertension who received liver transplantation.	**Inclusion criteria:** • Age ≥18 years • All patients with cirrhosis who were transplanted
	Exclusion criteria: • Patients who were transplanted due to acute liver failure. • Patients who are lost to follow-up in the first 5 years after liver transplant.
Clinical Use Case 4: Prediction of cardiac complications due to HCC treatments* **Note: includes surgical interventions, ablation, TACE, TARE, SIRT and immunotherapies*	**Inclusion criteria:** • Age ≥18 years • Diagnosis of HCC (any aetiology) • Cross sectional imaging follow-up (any modality) of liver diseases 6 months after treatment • Non-cirrhotic or no more than Child-Pugh B cirrhosis. • Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1 • Patients without history of prior HCC • Patients with a history of hypertension should be well controlled (< 140/90 mmHg) on a regimen of antihypertensive therapy. • With a minimum follow-up of two years or until death, after diagnosis of HCC
	Exclusion criteria: • Uncontrolled inter-current illness or psychiatric illness or social situations that would limit compliance with study requirements. • Subjects with history of another primary cancer • Fully recovered from any prior surgery and/or radiation and none within 2 weeks of initiating treatment. • Subjects with active hepatitis B or C on antiviral compounds may remain on such treatment, except for interferon.
Other populations (participation in control arms)	Subgroup of healthy controls **Inclusion criteria** : • Age ≥18 years • Subjects presenting cardiac fibrosis, without a known MASLD diagnosis (as controls for use case 2) **Exclusion criteria** : • Patients with diagnosis of MASLD

DETAILS ON SAMPLE SIZE AND POWER CALCULATION

1. Cohort studies:

Sample size calculation based on primary endpoints: The sample size calculation was calculated with a type I error of 5% and a power of 80%.

Clinical Use Case 1: Liver disease staging in MASLD patients

Clinical Coordinators: Pr Vlad Ratziu, and Dr Raluca Pais (ICAN), Dr. JM Pericas (VHIR)

A model to predict the probability of fibrosis changes is proposed, for both fibrosis progression and regression in MASLD patients. For the development of the model, inputs on the fibrosis changes observed in the cohort are needed. Fibrosis changes will be estimated from:

- at least one stage of fibrosis by liver biopsy
- an increase of 20-30% of baseline value by transient elastography
- clinical phenotype change from MASLD without significant fibrosis to MASLD with significant/advanced fibrosis; from MASLD to MASH

The model is also aimed to predict the risk of a rapid rate of fibrosis progression. The model's predictions may serve as continuous risk probabilities, discretized binary predictions (fast vs. slow progressors), or both. A rapid rate of progression is defined as either development of MASH or progression of liver fibrosis within one year or progression from no fibrosis to advanced fibrosis within 3 years (8).

Available evidence shows that around 20% of patients with MASLD develop MASH and amongst these, an average of 20% develop significant liver fibrosis (8-9). MASLD patients without baseline fibrosis progress 1 stage of fibrosis over 14.3 (95% CI, 9.1-50.0) years, whereas MASH patients do so on average within 7.1 (95% CI, 4.8-14.3) years (10). In addition, 33.6% patients are estimated to be progressors (increase in one stage of fibrosis) and 17-18% rapid progressors (from stage 0 fibrosis to stage 3 or 4 fibrosis- advanced- over a mean follow-up of 5.9, SD ±3.7 years, Singh et al Clin Gastroenterol Hepatol 2015). The main reference to test our hypothesis, namely that the combination of AI techniques and mechanistic models will be as efficacious as liver biopsy to detect the progression of liver fibrosis and to identify patients at risk of being rapid progressors, is that an average of 10% of patients will be rapid progressors within a 3-year timeframe. With these premises, and based on the average sample size of longitudinal studies that have investigated the rates of liver fibrosis progression in MASLD and MASH, **the estimated minimum sample size is as follows (Table 3): a minimum of 200 MASLD patients without MASH (80 with baseline biopsy) and 160 MASH patients (70 with baseline biopsy**) will be necessary to reach the necessary number of patients without either liver fibrosis or MASH and therefore detect progression. The expected number is much higher (see Table 6).

Clinical Use Case 2: MASLD and progression of cardiovascular diseases

Clinical Coordinators: Pr Norbert Frey and Dr Florian Leuschner (UKHD)

The average incidence of new cardiovascular disease in MASH patients is 8.3/1000 person-year, being higher in patients with advanced fibrosis (11). However, most large prospective cohorts have not included a composite endpoint of cardiovascular events that includes ischemic heart disease, heart failure, and atrial fibrillation, but rather have focused on the former and seldom also heart failure. Simon et al recently reported the difference in the incidence of major adverse cardiovascular events (MACEs, ie., ischaemic heart disease, stroke, congestive heart failure or cardiovascular mortality) between MASLD and non-MASLD patients to be 24.3 vs 16.0/1000 person-years; difference=8.3/1000 PY; aHR 1.63, 95% CI 1.56 to 1.70 (12). MASLD is associated with an average increase of 1.7-1.9 of new onset heart failure, which also increases along liver fibrosis (12). This equates to approximately an average of 27.4/1000-person year. In the most robust study to date, MASLD was only found to be associated with increased incidence of atrial fibrillation when significant fibrosis was present, with an average rate of 10.2 /1,000 person-year (13).

Consequently, for a combined endpoint with an estimated average incidence of 25/1000 persons-year, which would be almost twice as higher if focusing the analysis in the higher age ranges or patients with significant or **advanced fibrosis (of which the overall cohort will be enriched), the estimated sample**

size required is 1350 patients (Table 3), which is well under the projected sample size of several thousands.

The primary outcomes will be analyzed using Fine-Gray sub-distribution hazard model or Cox proportional hazards model. The Kaplan-Meier analysis will be performed to estimate and compare the risk of the outcomes in 3 years. The secondary outcome measures will be evaluated using multivariable logistic regression and/or Cox proportional hazard model. The model's prognostic performance will be evaluated using the omnibus test and AUROC analysis.

2. Exploratory analysis:

The statistical approach will be grounded on the use of Bayes Theorem, as to provide p-values that are more in line with confirmatory analyses (14)

Clinical Use Case 3: Patients with cirrhosis and Portal Hypertension receiving TIPS placement or liver transplantation

Clinical Coordinators: Pr Cristina Ripoll and Pr Alexander Zipprich (JUH)

MASLD is an emergent cause of decompensated cirrhosis, constituting one of its main causes, along with alcohol and viral hepatitis. Similar to MASLD, cirrhosis patients are at increased risk of heart failure. Few studies are available concerning the investigation of cardiovascular outcomes of decompensated MASLD/MASH cirrhosis and the placement of TIPS (15-17). However, Billey et al, (15) reported that hospitalization due to cardiac decompensation is observed in 20% of patients in the year after TIPS insertion. Accordingly, (18) asserts that the placement of TIPS in cirrhosis patients leads to a sudden increase in cardiac preload and output that can rapidly result in heart failure in the short term, or long-term cardiovascular changes, including cardiac volume overload. With the aim to improve risk stratification and patient selection, ARTEMIS will perform an exploratory analysis of the major cardiac complications related to TIPS placement. The total population size will need to be of **at least 120 patients (Table 3)**. Another treatment option for these patients that can lead to postoperative cardiac complications is liver transplantation (19-23). Both in the setting of liver transplantation and TIPS placement, portal hypertension and the characteristic systemic hemodynamic changes of cirrhosis play a role in the development of cardiac abnormalities, which can lead to cardiac decompensation in stress situations. As a consequence, ARTEMIs will perform for transplantation the same analysis as for TIPS. The expected sample size from the site's cohorts is much higher, with 1300 cases (Table 6).

Clinical Use Case 4: Prediction of cardiac complications due to HCC treatments

Clinical Coordinator: Dr Giuliana Amaddeo (AP-HP) and Dr Ivan Borbath (CUSL) and Dr Nicolas Lanthier (CUSL)

To date, heart complications associated to treatment addressed to HCC patients haven't proved to be significative in the scientific literature. In fact, it is estimated that around 10% of cases might present with syncope, whereas other complications, such pericarditis, arrythmias, heart failure, and myocardial infarction are only reported in clinical trials in an anecdotical fashion (24). However, with the onset of new immunotherapies and the combination thereof might pave the way to further investigations. For example, the understanding of the factors driving towards sensitivity or resistance to molecular target chemotherapy and liver resection remains poorly documented, as previous studies have mostly focused on conventional chemotherapy (25). Also, myocarditis cases have been reported as the most recurrent heart complication after the administration of combined immunotherapy, most notably, in the first weeks of administration (26). In terms of cardiotoxicity events related to the administration of immune checkpoint inhibitors (ICI), Shalata et al (27) asserts that considerable discrepancy has been observed among different trials; mostly due to misclassification and difficult diagnosis caused by ICIs, as well as the research conditions during the COVID pandemics. In addition, the exploratory analysis will also envisage to perform surrogate studies on the side-effects of immunotherapies on other organs and their impact on heart failure. In particular, the occurrence of pneumonitis due to the development of infiltrates in the interstitium and in the alveoli is of particular interest. In fact, this is the most recurrent immune-related adverse event in the lung, due to the administration of immunotherapies (28), Indeed, if not treated, pneumonitis can prove fatal (29). In an attempt to collaborate to the establishment of standard guidelines and facilitate the diagnosis of heart complications due to this therapeutic approach, ARTEMIs aims at conducting an exhaustive integrated exploratory analysis in the onset of cardiac complications related to therapeutic strategies addressed to

HCC patients, in different stages of the disease. For this exploratory analysis, the minimum size population should be of **at least 150 patients**, far below our expected sample size (Table 6)**.**

Table 3: Use Cases and respective statistical analysis plan.

Clinical Case 1 Progression of liver diseases in MASLD patients – Prediction model of fibrosis changes (progression and regression), with ability to distinguish between fast and non-fast fibrosis progression among MASLD patients	**Hypothesis**	The combination of AI and mechanistic techniques will be as efficacious as liver biopsy to detect changes between liver fibrosis phases, and to identify patients at risk of being rapid progressors.
	Minimum Sample Size	Minimum of 200 MASLD patients without MASH (80 with baseline biopsy) and 160 MASH patients (70 with baseline biopsy)
	Power calculation	Statistical power: 80% Type I error: 5%
Clinical Case 2 MASLD and progression of cardiovascular diseases	**Hypothesis**	The computational model will allow the evaluation of the prediction of cardiovascular diseases related to MASLD patients at 95% confidence interval.
	Minimum Sample Size	1350 patients
	Power calculation	Statistical power: 80% Type I error: 5%
Clinical Case 3 Cardiovascular complications in patients with cirrhosis and Portal Hypertension who receive TIPS placement or liver transplantation	**Hypothesis**	Hypothesis on cardiovascular complications in patients with cirrhosis after TIPS placement and transplantation will be derived from the exploratory analysis
	Minimum Sample Size	120 patients
	Power calculation	Based on Bayes Theorem
Clinical Case 4 Prediction of cardiac complications due to HCC treatments* **Note: includes surgical interventions, ablation, TACE and immunotherapies*	**Hypothesis**	Hypothesis on cardiotoxicity (myocarditis) related to HCC therapies will be derived from the exploratory analysis
	Minimum Sample Size	150 patients
	Power calculation	Based on Bayes Theorem

DESIGN OF THE CLINICAL STUDY

A multicenter, restrospective **longitudinal cohort of patients across the spectrum of liver disease.** Given the early maturity of virtual twins to support clinical practice, and the lack of awareness of physicians and patients on how these systems can positively impact on the clinical practice and the patients´ health outcomes, longitudinal cohorts are needed to accelerate the market and clinical uptake of these technologies. An observational non-interventional study is proposed.

The ARTEMIs cohort data is needed in the ARTEMIs R&D project for the development, refinement and performance and robustness validation of the computational models aimed to assist clinical decisions for a better management of patients along the MASLD spectrum.

Table 4 : Study design per Use Case

Study Design	**Clinical Case #1** Progression of liver diseases in MASLD patients	**Clinical Case #2** MASLD and progression of cardiovascular diseases	**Clinical Case #3** Cardiovascular complications in patients with cirrhosis and portal hypertension who receive TIPS placement or liver transplantation	**Clinical Case #4** Cardiovascular complications due to HCC treatments
Study type	Observational	Observational	Observational	Observational
Primary Purpose	Study of MASLD natural history	Evaluation of MASLD prognosis and cardiac comorbidities	Risk stratification on TIPS intervention and liver transplantation	Improve treatment administration to HCC patients at different stages
Minimum sample size required	200 patients	1350 patients	120 patients	150 patients
Observational model	Cohort study	Cohort study	Exploratory analysis	Exploratory analysis
Time perspective	Retrospective	Retrospective	Retrospective	Retrospective
Minimum follow-up period (months)	18-24 months	18-24 months	18-24 months	18-24 months

TYPE OF INTERVENTION

Only data recollection for their use in the training, testing and early validation of computational models (but no other intervention) will be performed.

DESCRIPTION AND TIMING OF STUDY PROCEDURES

Table 5 : Timeline for study milestones

Month 6	Final definition of the ARTEMIs cohort: final inclusion criteria, final selection of mandatory and optional variables for the recruitment of cases. Approval request to the Ethical Committees, if applicable. Deployment of local tools (Medexprim Suite™) started at first clinical sites.
Month 9	Readiness to start constituting the cohort and contracts signed with firsts data holder, most likely related to the incorporation of datasets from existing cohorts.
Month 13	Electronic Case Report Form (eCRF) finalized. Start of data collection from EHR, automated data source import where possible.
Month 18	Standardised and deidentified cases accessible for 3000 patients ready for use in computer models
Month 42	Start of evaluation of the utility of the Clinical Decision Support System (CDSS) on individual patients (on recent retrospective data, for which the outcome is known)
Month 48	End of the CDSS evaluation study. Path towards CE marking identified

2. PREPAREDNESS STATUS

DEVELOPMENT OF THE CLINICAL STUDY PROTOCOL

Scientific advice from regulatory and health technology assessment bodies

Betthera is a SME from the Czech Republic, participant in the ARTEMIS consortium and expert in regulatory affairs and assessing economic and social impact of medical technologies. Betthera will ensure that the system´s functionality meets the real needs of the European healthcare systems, that its complexity is proportionate to its potential added value, and will lead the description of the roadmap towards future CE marking and medical use of the system, following the MDR Regulation (EU) 2017/745).

Clinical efficacy, safety, and methodological guidelines (including guidelines on statistics)

The ARTEMIs cohort is an observational non interventional study. The therapeutic journey of the involved patients will not be affected by this study. There will be no interventions that affect the efficacy of therapy that the participant is routinely receiving.

Statistical analysis will be carried out by expert statisticians within the WP. An intermediate analysis will be carried out once data from half of the planned study cohort is available for both primary and secondary endpoints.

General principles of the statistical plan: All data will be collected in electronic format (e.g. csv, json, or other) in an anonymized form. In the analyses, a type I error of 5% will be considered and will be carried out with the Stata 15.1 program or higher. For each variable, a descriptive analysis will be carried out, expressing the qualitative variables as absolute number of cases and percentage of each of the categories in each cohort subgroup and overall. For the quantitative variables, the minimum and maximum means (standard deviations), medians (interquartile range) will be calculated. In all cases, the 95% confidence interval of the descriptive measures will be calculated. Patients with MASLD (Clinical Case 1) might differ in their background comorbidities and the pathways by which they enter the cohort in each site, thus leading different pre-test risk of having steatohepatitis (MASH) and liver fibrosis; also, they might have different baseline risk of developing cardiovascular complications (clinical case 2) depending on their cardiovascular risk factors and stage of fibrosis; the same applies to patients receiving TIPS (Clinical Case 3), particularly regarding the severity of cirrhosis in relationship with portal hypertension; moreover the likelihood MASLD patients of having HCC at baseline or developing HCC during follow-up (Clinical Case

4) depend upon certain baseline variables, remarkably having cirrhosis and certain comorbidities. Therefore, a comparison of the baseline characteristics of the sample subjects will be made to assess the differences within and across cohort subgroups.

Changes in categorical variables will be analyzed using a proportional probability logistic regression model appropriate for ordinal categorical data analysis, including treatment and adjustment for fibrosis stage and the presence of portal hypertension at baseline. Changes in continuous variables from baseline will be analyzed using an analysis of covariance (ANCOVA) with treatment as a factor adjusted based on fibrosis stage at baseline of the study. For time-to-event endpoints, treatment arms will be compared using the log rank test stratified by the above-mentioned stratification factors. The Hazard ratio with 95% CI will be determined based on a stratified Cox regression model to estimate the magnitude of the effect.

Involvement of citizens / patients, carers in drawing up the clinical study protocol

The European Liver Patient´s Association (ELPA) is a participant in the ARTEMIs project. Representatives of patients contribute to the definition of the cohort, and will supervise the cohort description to be incorporated in the deliverable D2.1, as part of the documental pack for the institutional ethics committees.

REGULATORY INTELLIGENCE

How the consortium will ensure access to regulatory expertise necessary to get advice on, and management of, regulatory affairs activities in all concerned jurisdictions?

Betthera will provide a regulatory roadmap needed for ARTEMIs regulatory acceptance. With their support, the consortium can rely on knowledgeable expertise and continuous strategic regulatory advice during the project lifetime.

How the consortium will ensure access to ethics expertise necessary to get advice on current proceedings and documentation requirements of all concerned ethics committees?

Based on the preliminary ethics assessment, the project coordinator, MAT, will coordinate the management of ethics considerations, in close cooperation with ethics committees of participating hospitals, and ELPA. Ethics considerations not only relate to the construction of the cohort and use of health data, but also to the development of a Clinical Decision Support System, including AI-based modules, and the recourse to animal models. With respect to the construction of the ARTEMIs cohort, each data provider partners (hospitals and research institutes) will involve their DPO to guide and supervise the procedure of ethics request approval. All data providers in this project have previous experience of participation in patient cohorts and clinical studies and are familiar with the elaboration of documentation packages for their institutional ethics committees. Some entities have units specialised in these matters which also offer them guidance and support.

The Project Technical Coordinator entity, Medexprim (MEDEX, member of the BCP group), will involve their legal department, with extensive experience in elaboration of contractual frameworks for international patient cohorts, elaboration of health data management plans (DMP) and data protection impact assessments (DPIA). MEDEX is responsible for the elaboration of the DMP and DPIA for ARTEMIs project, which will be part of the documentation package for each data provider.

SCIENTIFIC AND OPERATIONAL GOVERNANCE OF THE CLINICAL STUDY

Details about the sponsor(s)

The ARTEMIs Cohort will be coordinated by the Vall d'Hebron Research Institute (VHIR), in Spain

The cohort study is funded by the European Union Horizon Europe programme, under the call "HORIZON-HLTH-2023-TOOL-05-03 Integrated, multi-scale computational models of patient patho-physiology ('virtual twins') for personalised disease management", as part of the R&D Project "AceleRating the Translation of virtual twins towards a pErsonalised Management of fatty liver patients (ARTEMIs)".

The ARTEMIs **Cohort Coordinator** is Dr Raul Herance Camacho, responsible for the coordination of the participant sites as data providers.

The Principal Investigator at each data provider site is responsible for:

- ensuring legal and ethical compliance in the undertaken protocols, as well as compliance with quality standards.
- keeping copies of all versions of all site-specific documents, including protocols, and ethics approvals
- communication with their institution ethical committee and DPO for periodic updates on progress.

Composition, the role and the functioning of the planned board(s), governing bodies

The Governance bodies and roles are those of the ARTEMIs Project:

ARTEMIs Cohort Coordinator: Raul Herance (VHIR)

ARTEMIs Clinical Coordinator: Vlad Ratziu (ICAN)

ARTEMIs Scientific Coordinator: Irene Vignon-Clementel (INRIA)

ARTEMIs Technical coordination: Laure Saint-Aubert (Technical coordination team)

ARTEMIs Project coordination: Laura Muñoz, Amelia Suarez and Mario Aznar (coordination team).

ARTEMIs Project Coordination Board (PCB): It is responsible for supporting the project coordinator in the high- level managerial decisions, such as governing the consortium agreement, management of delays in achievement of milestones activation of contingency plan upon materialisation of risks resolution of disputes. The PCB is headed by the Coordinator and constituted by all PIs.

3. OPERATIONAL FEASIBILITY

HOW THE AVAILABILITY OF THE INTERVENTION(S) (INCLUDING COMPARATORS) IS SECURED THROUGHOUT THE ENTIRE IMPLEMENTATION PHASE

The consortium's partners (hospitals and research institutes) will provide access to health data for some 7500 cases (existing cohorts and routine care data) conservatively estimated as those cases which fulfil the inclusion criteria, including the minimum mandatory variables.

Recruitment targets per participant site: Data will be included in the ARTEMIs cohort using a common data model based on existing standards. Data will be deidentified before being made accessible to the partners.

Table 6: ARTEMIs cohort size per clinical case

Clinical Case #1&2 MASLD/MASH	Clinical Case #2 Cardiomyopathy subgroup	Clinical Case #3 TIPS / transplantation	Clinical Case #4 HCC	Controls	TOTAL
3850	650	1300 (650 TIPS, 650 Liver transplants)	1670	250	7720

Out of the total approx. 7500 cases, some will correspond to existing cohorts for which data is already de-identified and harmonised.

The projected cohort size for each category is well above the minimum sample required for each use case study (Table 4).

Table 7a: Brief description of the existing cohorts to be incorporated in ARTEMIs for **Clinical Case #1&2 MASLD/MASH**

Participant entity	Sample size	Clinical Case #1&2 MASLD/MASH: Cohort description
ICAN (France)	550	Single centre cohort of MASLD patients, longitudinal data for >2 years, approx. 80% with MASH, with extended metabolic profile + tissue biobank (ICAN´s contribution to LITMUS). Liver biopsy, serum and plasma samples for all.
VHIR (Spain)	400-1000	Around 400 MASLD/MASH patients with available retrospective data on liver biopsy and non-invasive tests of fibrosis. A proportion of these have also available serum/blood and stool samples. More than 1000 MASLD/MASH patients with available information on liver fibrosis assessed by non-invasive tests.
UKHD (Germany)	300	Hepatology Heidelberg (Prof. Michalski, Prof. Franck Billmann responsible for bariatric patients). Retrospective cases of bariatric surgery cases accessible with liver biopsy for all at the intervention (fresh frozen) and longitudinal serum analyses with information on medication (e.g. statins, SGLT2 inhibitors, beta-blockers). Liver MRI/CT scan infrequent. Cardiac examination infrequent, only if with previous cardiac pathology
HULAFE (Spain)	400	Liver Virtual Biopsy Project cohort of retrospective cases from the Valencia region, with longitudinal follow-up since 2017. Liver biopsy for all patients. Previous results: Extraction of radiomics features.
IMPERIAL (UK)	500	single centre cohort of MASLD patients, with liver biopsy and serum metabolomics and lipidomics profile for all, liver CT for some patients with advanced fibrosis.
CUSL (Belgium)	300	Single centre cohort of MASLD patients with laboratory data, anthropometric data, medications. Follow-up with hepatic elasticity and cardiovascular events. Liver biopsy for some of them. No systematic CT or MRI. Among them, a subgroup with systematic hepatic MRI and biopsy + frozen liver and frozen serum and characterization also on the muscle compartment. Among the above cohort, a percentage of patients have MASLD complicated by cirrhosis with clinical and biological follow-up (most frequently followed by ultrasound imaging).
CHA (Germany)	200	[pending description]
ULS (Italy)	300	Single center cohort of patients with different metabolic disorders (MASLD, DMT2). In this cohort, liver MR elastography, iron concentration, fat fraction, CT/MRI body composition, laboratory data, anthropometric data, and medication are available A subgroup 120 patients includes CT coronary angiography.
MUV (Austria)	500	Single center cohort with 500 MR data sets for chronic liver disease patients at different stages and of different aetiologies including MASLD/MASH. Some of this cohort has liver MR elastography, iron concentration, fat fraction, laboratory data, biopsy, anthropometric data, and Information on concomitant medication.

ULEI (Germany)	400	Large biobank with linked clinical data for chronic liver disease patients at different stages and of different aetiologies. Cases available for MASLD/MASH aetiology with serum and fresh frozen tissue for some of them, MRI or CT scan available. Information on concomitant medication (e.g. statins, beta-blockers, SGLT2 inhibitors)
TOTAL	**3850**	

Table 7b: Brief description of the existing cohorts to be incorporated in ARTEMIs for **Clinical Case #2 Cardiovascular disease subgroup**

Participant entity	Sample size	Clinical Case #2 Cardiomyopathy subgroup: Cohort description
VHIR (Spain)	300	2 single centre cohorts of the liver-heart-brain axis in 100 + 200 T2DM patients, with 5 years follow-up. Available data includes sarcopenia, coronary artery calcification, epicardial adipose tissue, echo-cardio, myocardial insulin resistance. Previous results: Phenotyping of T2DM patients in terms of insulin resistance.[1] Moreover, a substantial proportion of the patients in the MASLD fibrosis cohort for CC1 have data available on echocardiography and other cardiac tests.
UKHD (Germany)	350	Cardiology Heidelberg (Dr Leuschner/Dr Wienecke). Cases of cardiac fibrosis with clinical data and samples for some of them (cardiac fresh frozen tissue and serum), with a minimum follow up of 1 year accessible. Cases of other cardiology patients with diagnosed MASLD/MASH, clinical data and samples for some of them (cardiac fresh frozen tissue and serum), with a minimum follow up of 1 year accessible.
TOTAL	**650**	

Table 7c: Brief description of the existing cohorts to be incorporated in ARTEMIs for **Clinical Case #3 TIPS / transplantation**

Participant entity	Sample size	Clinical Case #3 TIPS / transplantation: Cohort description
VHIR (Spain)	50	Available data on retrospective cases of TIPS. Not all of them have available information on echocardiography.
AP-HP, Pitié Salpêtrière (France)	200	Single centre cohort of 200 cirrhosis patients with TIPS placement, longitudinal clinical, biological and morphological data for before TIPS to at least 1 year follow-up, HVPG before and immediately after intervention. Cardiac echography for all before intervention, and after intervention only for those who developed cardiac decompensation.
AP-HP, Paul Brousse (France)	300	Liver transplant patients with longitudinal data including serum samples, CT/MRI scan, liver biopsy every 1-2 years and cardiac examination (cardiovascular complications is the main cause of death in this intervention.
JUH (Germany)	100	TIPS cohort with clinical data (age, gender, comorbidities, arterial, venous portal venous pressure, incidence of non-hepatic events), medical imaging data (CT – with contrast agent (4 phases: native, arterial, portal venous, venous); MRI

		– with contrast agent (Gadoxetic acid: in-/opp phase, native T1, DCE, DWI, T2_haste, T2_blade, hepatobiliary phase) + CT (partially GSI). Ultrasound (HVPG). Routine blood tests with liver and renal function, lipid profile, Inflammation markers. Biopsies very rarely.
HULAFE	400	A single-center cohort comprising cirrhosis patients who underwent TIPS placement (n=140) and/or liver transplant (50-60/year), with clinical and biological data (such as hemodialysis prior to TIPS, age, sex, MELD-Na score, MELD 3.0 score, Child-Pugh score, and presence of portal thrombosis/cavernomatosis) and medical imaging data (CT or MRI scans) performed 1-3 months before TIPS placement, with follow-up extending to at least 2 years.
CUSL (Belgium)	100	Prospective single-centre list of patients evaluated for liver transplantation (n=200) or TIPS (n=20) with clinical and body composition data (CT-scan). Some of these patients received a transplant (n=100) or TIPS (n=6). Re-evaluation of histology and body composition at 6 months post-intervention.
MUV (Austria)	150	Liver transplant patients with longitudinal data including Gadoxetic acid enhanced-MRI and MRCP, including Functional Liver Imaging Score (FLIS) and laboratory data, anthropometric data, and medication. Survival data including rejection, re-transplantation or death. Basic cardiac examination and very rarely Biopsies.
TOTAL	**1300**	

Table 7d: Brief description of the existing cohorts to be incorporated in ARTEMIs for **Clinical Case #4 HCC**

Participant entity	**Sample size**	**Clinical Case #4 HCC: Cohort description**
VHIR (Spain)	120	Patients that have received immunotherapy on HCC clinical trials. Only 2-3 cases of suspected cardiovascular events due to immunotherapy.
HULAFE (Spain)	200	Retrospective cases diagnosed of HCC and treated with surgical intervention, locoregional therapies (ablation, TARE, TACE) and/or immunotherapies with follow-up imaging (MRI or CT) ≥ 6 months after the treatment.
AP-HP, Henri Mondor (France)	300	Resected HCC histological slides and respective clinical data (PATHEP cohort)
JUH (Germany)	300	[pending description]
CUSL (Belgium)	350	Retrospective single-centre cohort of patients with HCC (all causes) (n=250) and prospective cohort of patients undergoing immunotherapy for HCC (all causes) (without a 2-year follow-up at the moment) (n=100). Consideration of treatment-related complications.
CHA (Germany)	200	[pending description]
ULEI (Germany)	200	Retrospective cases diagnosed of HCC and treated with surgical intervention or transplantation, partly with imaging (MRI or CT) before the treatment.
TOTAL	**1670**	

Table 7e: Brief description of the existing cohorts to be incorporated in ARTEMIs for **Control cases**

Participant entity	Sample size	Control cases: Cohort description
HULAFE (Spain)	100	Retrospective cases diagnosed with lung cancer, colorectal cancer, breast or prostate patients, with no liver condition during their entire follow-up.
IMPERIAL (UK)	50	Endometrial cancer patients, with no liver condition during their entire follow-up
MUV (Austria)	100	Retrospective more than 100 cases with suspected diagnosis of benign and malignant hepatobiliary diseases, but with no liver condition during their entire follow-up.
TOTAL	**250**	

HOW THE STUDY POPULATION WILL BE RECRUITED

Patients along the MASLD spectrum will be recruited in 12 participant sites. A researcher will select cases fulfilling the inclusion and exclusion criteria and process them according to the Project work plan.

The recruitment rate will be monitored at each site, by periodical review of the screening log and internal meetings to identify unexpected losses of inclusion. The project coordinator will globally monitor the recruitment evolution of the cohort.

In case of slow/insufficient recruitment, the following mitigation strategies are planned:

- Revision of the inclusion/exclusion criteria
- Positive incentives for recruitment
- Evaluation of additional clinical sites

How many clinical sites will contribute to the recruitment of the study population in which countries? Are these clinical sites part of an established clinical trial network? Please also describe the selection criteria of the clinical sites.

A total of 12 university hospitals as well as medical research centres linked to hospitals with high-volume admission of chronic liver disease patients, and with demonstrated experience conducting clinical trials, will contribute to the recruitment, as summarised in Table 7a-e.

Table 8: Recruitment centres for ARTEMIs cohort

	Clinical site, Short name	Clinical site	Country
1	APHP	AP-HP (Henri Mondor, Paul Brousse and Pitié Salpêtrière)	France
2	JUH	Jena University Hospital	Germany
3	VHIR	Vall d´Hebron research institute	Spain
4	UKHD	Heidelberg University Hospital	Germany
5	CUSL	Cliniques universitaires Saint-Luc	Belgium
6	HULAFE	La Fe Health Research Institute	Spain
7	CHA	Charité	Germany
8	ULS	Sant'Andrea University Hospital, Sapienza University of Rome	Italy
9	IMPERIAL	Imperial College healthcare, NHS trust	UK

10	MUV	Medical University of Vienna	Austria
11	ULEI	Leipzig University Hospital	Germany
12	ICAN	Institute of Cardiometabolism and Nutrition	France

In the **12** sites, research structures are already established, and the respective clinical leaders have experience on coordination of national and international clinical trials. The sites have been also selected to obtain a sample of patients from 7 different countries, avoiding biases of too homogeneous samples.

Non requirement of informed consent from patients:

It is expected that both the institutional Ethics Committee of participant sites in the ARTEMIS cohort study will accept the non-requirement of informed consent from patients based on the International Ethical Guidelines for Health-Related Research Involving Human Subjects prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO) (2016) which state that:

" A research ethics committee may approve a modification or waiver of informed consent

to research if:

· the research would not be feasible or practicable to carry out without the

waiver or modification;

· the research has important social value; and

· the research poses no more than minimal risks to participants. "

It is therefore expected that the ARTEMIS cohort participant sites will be exempted from requesting informed consent from the patient given that:

- There may be difficulties in obtaining informed consent from patients who are already deceased.

- The project has an important societal value as it will result in a virtual twin allowing a better understanding of the underlying mechanisms involved in MASLD progression, as well as the prediction of cardiovascular events at different stages of the disease

- The research involves only in-silico data handling and analysis, with no interventional activity and no risk to participants.

Will recruitment of the study population be of competitive nature between the clinical sites? (Please describe how underperformance of individual clinical sites in recruitment will be managed.)

No. All clinical sites will compromise a minimum recruitment of patients through the project.

What evidence supports the ability of the individual clinical sites to recruit the required number of study participants within the planned timeline (e.g. documented performance in previous clinical studies of similar complexity targeting very similar study population)?

Table 9: Recruitment capability of the centres participant in the ARTEMIs cohort

	Clinical site	Justification of their ability to recruit
1	AP-HP (Henri Mondor, Paul Brousse and Pitié Salpêtrière)	AP-HP comprises 39 hospitals treating 7 million patients a year for consultations and emergencies, scheduled hospitalisations or at-home care. In ARTEMIs project, APHP is represented by 3 hospitals:

		Henri Mondor, Paul Brousse and Pitié Salpêtrière
2	Jena University Hospital	53,000 in- and 275,000 outpatients per year The Department of Gastroenterology, Hepatology and Infectiology at the University Hospital Jena offers the full range of diagnostics and treatments for diseases including liver.
3	Vall d´Hebron research institute	Clinical trials (1400 on-going in different medical areas) National and European projects Expertise in: – MASLD/MASH, HCC, Cardiology, Bariatric surgery, Endocrinology, Radiology, Nuclear Medicine – Medical image analysis – Metabolomics/lipidomics – Clinical and Preclinical studies (Liver-Heart) (With and without imaging)
4	Heidelberg University Hospital	With approximately 800,000 outpatients and 100,000 inpatients treated each year, the University Hospital is one of the ten largest hospitals in Germany. The Department of Cardiology, Angiology, and Pneumology at Heidelberg University Hospital is one of the leading university cardiology departments in Germany, with about 12,500 inpatients with cardiovascular diseases, treated per year.
5	Cliniques universitaires Saint-Luc	One of the main academic hospitals of Brussels, The Hepato-gastroenterology Department and the Gastroenterology and Hepatology Center (GAEN) carry out numerous fundamental research projects, translational and clinical research in various fields of the discipline, including studies of new drug treatments for MASH and studies of the links between liver disease and insulin resistance
6	La Fe Health Research Institute	230,000 inpatients +outpatients treated each year. The hepatobiliopancreatic surgery and transplant unit of this centre has performed more than 3,000 liver transplants since it started the liver transplant programme in 1991. This makes it the first centre in Spain to reach this figure.
7	Charité	Every year, 130,000 patients undergo inpatient treatment at the clinic, and about 1,000,000 are treated in the outpatient clinics. The Center for Diagnostic and Interventional Radiology and Nuclear Medicine includes the interventional radiology section that provides targeted, minimally invasive local therapies for a variety of diseases including local cancer treatment (e.g., ablation of liver tumors). The Department of Hepatology and Gastroenterology, as a center of maximum care, deals with the diagnosis and treatment of diseases of the liver and biliary tract, the pancreas and the gastrointestinal tract, in close cooperation with other specialist disciplines such as surgery, radiology or nuclear medicine.
8	Sant'Andrea University Hospital, Sapienza University of Rome	The Sant'Andrea Hospital is a teaching hospital, which houses the Faculty of Medicine at the "Sapienza" University of Rome
	Imperial College healthcare, NHS trust	Our liver unit is based at St Mary's Hospital, with weekly clinics located at Charing Cross, Hammersmith and St Charles hospitals.

9		St Mary's Hospital has 246 beds and handles 22,685 admissions each year. The liver unit provides a range of services to treat the broad spectrum of conditions affecting the liver, including cirrhosis,liver cancer and fatty liver disease. Our day unit at St Mary's Hospital provides a day-case service for procedures such as liver biopsy, paracentesis, venesection and fibroscan assessment. We also perform hepatic encephalopathy testing (PHES, MRI) on patients with cirrhosis.
10	Medical University of Vienna	61,016 patients treated as inpatients each year 515,687 patients treated as outpatients each year MUV is the largest medical training centre in the German-speaking countries. MUV Clinical Department of Gastroenterology and Hepatology deals with the liver diseases. Gastroenterological and hepatological diseases, the various aspects of nutrition, metabolism and psychosomatics are our focus of interest..
11	Leipzig University Hospital	52,527 inpatient and day patient cases (2021) 305,642 outpatients receiving treatment (2021) Clinic and Polyclinic for Oncology, Gastroenterology, Hepatology, Pneumology and Infectious Diseases The Hepatology department offers the most modern examination methods and treatment options for patients with liver diseases. Outpatient treatment takes place in the hepatological outpatient clinic, the liver cancer center and the liver transplant outpatient clinic.The Hepatology department is actively involved in conducting national and international clinical studies to improve the diagnosis and treatment options for patients with diseases of the liver and bile ducts.
12	Institute of Cardiometabolism and Nutrition	ICAN participates on several research projects against MASH: – EPOS / LITMUS – HOTSURFER – CORONASH for natural history and diagnostic tools – EU-PEARL for processing. The AP-HP and ICAN teams created the MASH clinic to offer patients an innovative, multidisciplinary and personalized care pathway. This course aims to anticipate and intercept the complications of MASH (early atherosclerosis, arterial hypertension, diabetes, etc.) and to offer personalized care to each patient. The MASH clinic is the first hospital structure for the diagnosis and multidisciplinary management of patients with metabolic steatosis in France.

ADDITIONAL SUPPLY

MEDEX will deploy its processing tools at the hospitals. This includes an eCRF which will be tailored to the project and that users will use to collect all retrospective data. The solution will be connected to the hospital's PACS and to other data sources where possible, to automate as much as possible the data collection and de-identification process. If the site already has some data collection in a specific format (eg. excel), data will be mapped to the target format.

For the secondary analysis of liver biopsy tissue samples, mass spectrometry-based proteomic analysis will be performed on a subset of samples already available in the biobanks of existing cohorts, in order to examine: cell surface receptors and ligands, the extracellular matrix (ECM) proteins, secreted soluble factors as well as global proteome of liver tissue and serum/plasma samples. Tissue samples are pulverised and sonicated before protein clean-up and digestion, using the SP3-protocol that has been migrated to a Bravo liquid handling robot for automated 96-plex sample processing, and will be further optimised for low-input applications by reducing sample volumes. For increased sensitivity and accurate absolute quantification of key proteins, we will use PRM-based proteomics in combination with spike-in of heavy stable isotope-labeled calibrator proteins. This approach has been established in the Klingmüller group at the participant entity DKFZ, for several transmembrane receptors and signal transduction proteins and can be extended.

For a subset of biological samples also available in the biobanks of existing cohorts, existing metabolomics and lipidomics pipeline analyses will be performed at VHIR, using a Bruker Avance || 600 MHz spectrometer with the Carr–Purcell–Meiboom–Gill (CPMG) spin-echo pulse sequence. Serum samples will be received at -80° and prepared according to standard protocol, 1H-NMR spectra will be acquired and binned into variable sized packets using the MestreNova software and stored into the Simca software. For the metabolomic and lipidomic analyses of liver tissue samples, polar and non-polar metabolites will be extracted from the tissue and then a similar procedure to the analysis of biofluid samples will be followed.

PLANS ON DATA MANAGEMENT ASPECTS

The detailed Data Management Plan (DMP) of ARTEMIs will be delivered by MEDEX (D2.1, "Initial Data Management Plan, including a detailed description of the expected ARTEMIs cohort (T2.2), a review of the applicable regulatory framework (T1.4), a DPIA and Data access/sharing agreement models (T2.3)" led by MEDEX and due in month 6. The DMP is updated at each reporting period (M18 and M36).

The DMP elaboration will be greatly facilitated by the experience already developed in their commercial projects with the pharmaceutical industry, as well as in R&D projects such as PRIMAGE, CHAIMELEON and EUCAIM in the field of multicentre and multimodal health data management. The DMP elaboration and updates are part of WP2. MEDEX will be responsible for assessing risks in data protection, developing the contractual framework between partners related to access to data and ensuring FAIR principles are respected on the ARTEMIs cohort. INRIA will ensure that research and experimental data, and more generally all research outputs, also follow the same principles, adhering to guidelines such as the "Guidelines for constructing, verifying and validating models" (ISO/DTS 9491-1). Below is a summary of how the data/research outputs will be managed in line with the FAIR principles:

- Types of data/research outputs: This project will primarily reuse pre-existing data corresponding to around retrospective patient data (coming from existing collections or from routine practice) and preclinical data (animal models, in-vitro experiments). Data will be organised in training, testing and validation datasets. The project will also generate data from processing pipelines (omics) and experiments.
- Findability of data/research outputs: The ARTEMIs cohort will be composed of different data collections uniquely identified and contributed by partners. A CDM will be defined to ensure data is easily findable and specific datasets can be created from the global ARTEMIs cohort. A persistent and unique identifier will be attributed to each dataset, with proper versioning management. Similar principles will be applied to submodels and models, software and experimental data.
- Accessibility/reusability of data/research outputs: The consortium will apply the "as open as possible, as close as necessary" principles, ensuring a fair balance between large dissemination and data protection and commercial exploitation. Issues related to the ownership of intellectual property rights will be framed within a specific contract between the partners that will specify any limitations on use,

particularly in terms of purpose, exclusivity or not, duration, geographical area and compensation. Results will be made public whenever possible and only in exceptional cases they will remain private either for an embargo period only, or permanently, and this will be determined on a per-result basis. An open access archive for selected project outcomes will be designated (e.g. Zenodo).

- Interoperability of data/research outputs: Standard ontologies and vocabularies/formats will be used for the data, including HL7-FHIR, DICOM, SNOMED-CT, LOINC. Model developers will also use interoperability standards such as Systems Biology Markup Language (SBML).
- Curation and storage/preservation costs: Curation of health datasets is foreseen to ensure quality of data for the development of the computational models. Peer-review internal processes will be set-up for deliverable reports. Each data holder will be responsible for storing their data, preserving them and making them accessible to the federated infrastructure according to MEDEX's specifications. MEDEX will be responsible for the central components of the architecture, in particular the security aspects. For the project generated data, including deliverables, the Coordinator will provide a private archive accessible only to consortium members (Project Repository) for project management resources that needs to be shared among all participants.

Data collection and management including analysis, reporting, security:

The data management protocols for this cohort will be approved by the institutional DPO at each participant clinical site. The P.I. or his/her designees must ensure that data are correctly recorded in the designated databases stored at the participating centre´s local IT systems (the Medexprim Suite™ solution), and completely by authorized personnel. The investigator will confirm the integrity of the data transferred to the designated databases by signature.

All completed study related documents must be archived by each Principal Investigator at least for the minimum legal period as established in their ruling national laws for clinical studies.

The P.I. is responsible for the complete de-identification of the patient data at the institution´s premises before making the datasets available to other Project Partners. Data de-identification will be done in full compliance with the national laws for GDPR enforcement. All necessary mechanisms to ensure data quality and integrity will be in place. MEDEX will provide solutions to support these activities.

The consortium partners may sign Data Transfer Agreements and make their data accessible in a central storage hub to be deployed by MEDEX, or alternative make their data accessible at their premises for distributed training of the models, providing access to the required computational resources.

Data analysis and reporting are part of the activities involved in the elaboration of the Cohort, and thus contemplated in the Work plan.

The funding for this cohort is requested to the European Commission via the Horizon Europe programme for research funding. The participants in ARTEMIs project are committed to contribute to the overall imperative of Open Science. As part of the Project, the Data Management plan to be delivered in month 6, will also contemplate the open access to some specific datasets. The mechanisms to enable access to deidentified datasets in full compliance with GDPR and national laws will be elucidated during the Project. Authorisation to the involved IEC and DPO will be requested, as needed.

Data reuse in other research projects

The ARTEMIs cohort may be reused in future R&D projects. The ARTEMIs cohort coordinator would inform each cohort participant centre of any relevant request for access to the ARTEMIs cohort data, and request their written approval. Any re-use shall comply with the national law of each partner and in all cases with European laws, including but not limited to the GPRD, and shall comply with the administrative requirements of each partner. If necessary and applicable, any re-use must be the subject of a submission to the local ethics committee and a duly signed contract.

REPORTING OBLIGATIONS TO REGULATORY BODIES AND ETHICS COMMITTEES

The ARTEMIs cohort is a multicentric cohort involving 12 university hospitals and medical research centres in 7 countries.

The PI for each clinical site will strictly follow the recommendations and obligations for reporting of progress and results, as established by their institution´s IEC and institutional DPO.

The Project Coordinator, the Cohort Coordinator and the PI at each clinical site, are aware that an Audit Inspection can be carried out by the institution, to confirm that the study is conducted as per protocol, In case of an audit, the investigators will permit a direct access to all study documents, accountability records, and source data.

RESPONSABILITIES HELD BY EXTERNAL ENTITIES

No external entities will support the ARTEMIs consortium in the constitution of the cohort, all activities are internalised.

PLANS FOR MAJOR STUDY MILESTONES AND EVIDENCE SUPPORTING ITS FEASIBILITY

Table 10 : Main milestones of the cohort study

Milestones of the ARTEMIs cohort	Plan
Milestone 1: Compiling the required regulatory and ethics submission package	This work will be done from project start to month 9 under the coordination of Medexprim, as part of T1.5 (“Ethics requirements) and T2.3 (“Data protection assessment and contractual”). Each participant clinical site is responsible for preparing the study protocol documentation for their IEC and DPO. Medexprim is committed to give full support for documentation of the data management plan and impact assessments.
Milestone 2: Receipt of regulatory and ethics approval	This task will initially focus on existing cohorts made accessible at participant sites such as ICAN, APHP, VHIR, HULAFE or CUSL, to name some. Approval for access to these collections is expected in month 9. Additionally, clinical site will also incorporate cases from its EHR corresponding to patients routinely attended (approval expected by month 12), until the total recruitment target is met.
Milestone 3: Initiation of clinical site	All clinical sites will work together with Medexprim in WP2 to deliver a common data model based on existing standards. As part of T2.7, data collection and interoperability solutions will be deployed by Medexprim at all sites to facilitate the creation of a harmonised cohort.
Milestone 4: 3000 cases incorporated in the ARTEMIs cohort	Clinical partners and modellers will iterate about the necessary quality of the data for modelling of all kinds, to reach a consensus of minimum data variables and standardised format needed per use case. As part of T2.8, under the coordination of VHIR, the data will be made accessible for project developments, linked to milestone MS2 “ARTEMIs cohort available” due in month 18 and corresponding to 3,000 cases incorporated.
Milestone 5: Completion of the cohort and reporting of the study results.	The cohort is completed, with 7,500 cases incorporated, which have been used for the development and evaluation of the computational models, the assessment of the results is linked to D2.4 “Final DMP, including a description of the final cohort” due in month 48.

4. BIBLIOGRAPHIC REFERENCES

1. Eslam M, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M, et al. A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. J Hepatol. juill 2020;73(1):202-9.

2. Yuan Q, Wang H, Gao P, Chen W, Lv M, Bai S, et al. Prevalence and Risk Factors of Metabolic-Associated Fatty Liver Disease among 73,566 Individuals in Beijing, China. Int J Environ Res Public Health. 13 févr 2022;19(4):2096.

3. Pais R, Maurel T. Natural History of NAFLD. J Clin Med. 10 mars 2021;10(6):1161.

4. Ekstedt M, Nasr P, Kechagias S. Natural History of NAFLD/NASH. Curr Hepatol Rep. 2017;16(4):391-7.

5. Tsochatzis EA. Natural history of NAFLD: knowns and unknowns. Nat Rev Gastroenterol Hepatol. mars 2022;19(3):151-2.

6. Cvitanović T, Reichert MC, Moškon M, Mraz M, Lammert F, Rozman D. Large-scale computational models of liver metabolism: How far from the clinics? Hepatology. 2017;66(4):1323-34.

7. Rinella ME. Nonalcoholic Fatty Liver Disease: A Systematic Review. JAMA. 9 juin 2015;313(22):2263.

8. Singh SP, Misra B, Kar SK, Panigrahi MK, Misra D, Bhuyan P, et al. Nonalcoholic fatty liver disease (NAFLD) without insulin resistance: Is it different? Clin Res Hepatol Gastroenterol. 1 sept 2015;39(4):482-8.

9. Allen AM., Therneau TM., Ahmed OT., Gidener T., Mara KC., Larson JJ., et al. Clinical course of non-alcoholic fatty liver disease and the implications for clinical trial design. J Hepatol. 2022 Nov;77(5):1237-1245.

10. Sanyal AJ, Van Natta ML, Clark J, Neuschwander-Tetri BA, Diehl A, Dasarathy S, et al. Prospective Study of Outcomes in Adults with Nonalcoholic Fatty Liver Disease. N Engl J Med. 21 oct 2021;385(17):1559-69.

11. Mantovani F, Clavel MA, Michelena HI, Suri RM, Schaff HV, Enriquez -Sarano Maurice. Comprehensive Imaging in Women With Organic Mitral Regurgitation. JACC Cardiovasc Imaging. avr 2016;9(4):388-96.

12. Simon TG, Roelstraete B., Alkhouri N., Hagström H., Sundström J., Ludvigsson JF. Cardiovascular disease risk in paediatric and young adult non-alcoholic fatty liver disease. Gut. 2023 Mar;72(3):573-580.

13. Kleef LA van, Kavousi M, Knegt RJ de. Reply to: “Liver stiffness, fatty liver disease and atrial fibrillation in the Rotterdam study: Some issues”. J Hepatol. 1 nov 2022;77(5):1467-8.

14. Moyé L. What Can We Do About Exploratory Analyses in Clinical Trials? Contemp Clin Trials. nov 2015;45(0 0):302-10.

15. Billey C, Billet S, Robic MA, Cognet T, Guillaume M, Vinel JP, et al. A Prospective Study Identifying Predictive Factors of Cardiac Decompensation After Transjugular Intrahepatic Portosystemic Shunt: The Toulouse Algorithm. Hepatology. déc 2019;70(6):1928.

16. Ali A., Sarwar A., Patwardhan V.R., Fraiche A.M., Tahir M.M., Luo M., et al. Echocardiographic and Other Preprocedural Predictors of Heart Failure After TIPS Placement in Patients With Cirrhosis: A Single-Center 15-Year Analysis. AJR Am J Roentgenol. 2022 Jul;219(1):110-118.

17. Modha K., Kappoor B., Lopez R., Sands M.J., Carey W. Symptomatic Heart Failure After Transjugular Intrahepatic Portosystemic Shunt Placement: Incidence, Outcomes, and Predictors. Cardiovasc Intervent Radiol. 2018 Apr;41(4):564-571.

18. Rajesh S, George T, Philips CA, Ahamed R, Kumbar S, Mohan N, et al. Transjugular intrahepatic portosystemic shunt in cirrhosis: An exhaustive critical update. World J Gastroenterol. 7 oct 2020;26(37):5561-96.

19. Sonny A., Govindarajan S.R., Jaber W.A, Cywinski J.B. Systolic heart failure after liver transplantation: Incidence, predictors, and outcome. Clin Transplant. 2018 Mar;32(3):e13199.

20. Mohamaddi F., Ramachandran J., Woodman R., Muller K., John L., Chen J., et al. Impact of cardiac dysfunction on morbidity and mortality in liver transplant candidates. Clin Transplant. 2022 Jul;36(7):e14682.

21. Snowden C.P., Hughes T., Rose J., Roberts D.R. Pulmonary edema in patients after liver transplantation. Liver Transpl. 2000 Jul;6(4):466-70.

22. Qureshi W., Mittal C., Ahmad U., Alihayim Z., Hassan S., Qureshi S. et al. Clinical predictors of post-liver transplant new-onset heart failure. Liver Transpl. 2013 Jul;19(7):701-10.

23. Saad A.A., Arman H.E., Shamseddeen H., Elsner N., Elsemesmani H., Johnson S. et al. Cirrhotic cardiomyopathy: Predictors of major adverse cardiac events and assessment of reversibility after liver transplant. J Cardiol. 2023 Aug;82(2):113-121.

24. Llovet JM, Castet F, Heikenwalder M, Maini MK, Mazzaferro V, Pinato DJ, et al. Immunotherapies for hepatocellular carcinoma. Nat Rev Clin Oncol. mars 2022;19(3):151-72.

25. Liu KL, Chen JS, Chen SC, Chu PH. Cardiovascular Toxicity of Molecular Targeted Therapy in Cancer Patients: A Double-Edged Sword. Acta Cardiol Sin. juill 2013;29(4):295-303.

26. Chung WB, Youn JC, Youn HJ. Cardiovascular Complications of Novel Anti-Cancer Immunotherapy: Old Problems from New Agents? Korean Circ J. 27 mai 2020;50(9):743-53.

27. Shalata W, Abu-salman A, Steckbeck R, Mathew Jacob B, Massalha I, Yakobson A. Cardiac Toxicity Associated with Immune Checkpoint Inhibitors: A Systematic Review. Cancers. 18 oct 2021;13(20):5218.

28. Porcu M, De Silva P, Solinas C, Battaglia A, Schena M, Scartozzi M, et al. Immunotherapy Associated Pulmonary Toxicity: Biology Behind Clinical and Radiological Features. Cancers. 5 mars 2019;11(3):305.

29. Cui T ming, Liu Y, Wang J bei, Liu L xin. Adverse Effects of Immune-Checkpoint Inhibitors in Hepatocellular Carcinoma. OncoTargets Ther. 16 nov 2020;13:11725-40.