ARTEMIS

The Power of Virtual Twins to Fight MASLD

D9.3 – PRELIMINARY KER EXPLOITATION AND DATA SUSTAINABILITY STRATEGY

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.	D9.3
Title	Preliminary KER exploitation and data sustainability strategy
Work Package	WP9
Dissemination level	PU – Public
Nature	R – Document, report
Responsible partner	Matical Innovation (MAT)
Main Contact person	Pedro Salazar
Deliverable due submission date	31/12/2025
Deliverable actual submission date	19/12/2025

Document validation:

Validated by	Name	Organisation short name	Visa
Responsible	Amelia Suarez	MAT	OK
Reviewer 1	Petr Straka	BETT	OK
Reviewer 2	Ehud Zigmond	SHEBA	OK

Disclaimer

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.

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 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.

1 EXECUTIVE SUMMARY

This deliverable presents the preliminary exploitation and data sustainability strategy of the ARTEMIS project. It is structured around the Key Exploitable Results (KERs) that are expected from the developments described in the DOA. As an interim plan, it reflects the results of a scouting and positioning exercise conducted during the first half of the project, in which initial value propositions, potential exploitation pathways and enablers and barriers were identified for the most promising assets.

Based on the list of KER documented in the DOA, this report evaluates the current maturity, innovation potential and exploitability of the KERs, with a detailed focus on the three most actionable results at present (KER1, KER2 and KER3). These analyses provide preliminary strategic guidance on exploiting multimodal data processing pipelines, clustering tools and Virtual Twin (VT) models. However, developments are ongoing and the final scope of achievements and the exact contributions of partners are yet to be confirmed as the results reach higher development maturity.

In the second half of the project, these interim exploitation and sustainability plans will be carefully reviewed with the consortium, informed by stronger evidence from model validation activities, clinical integration work, and stakeholder feedback. This will enable the final exploitation strategy to select the most impactful assets, describe their selected IPR protection mechanism and define their business models. Also, to establish concrete measures for longer-term access and sustainability.

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 INTRODUCTION

2.1 Background

The deliverable, D9.3, is the mid-term analysis of how ARTEMIS Key Enabling Results (KERs) can be exploited and how access to data and models can be sustained after the project ends. This report builds on the study plan defined in D9.1 and the value propositions outlined in D9.2 It prioritises three KERs that are currently at a higher development stage and have a relevant degree of innovation and exploitability. A detailed functional, technology, and market analysis has been carried out for these KERs. The interim sustainability plan describes the sustainability principles and actions that need to be agreed and implemented before the end of the project, and identifies the key partners involved in the data provision and model development. Since data gathering, models' development and other advancements towards the delivery of the KERs are on-going activities, the analyses provided here should be considered preliminary and intended to guide the refinement of exploitation and sustainability decisions to be made in the second half of the project.

2.2 Objectives

D9.3 is one of the deliverables of the WP9 entitled "Exploitation and Sustainability" and lead by MATICAL. WP9 has as general objective to contribute to the Project Objective 7: To accelerate the clinical uptake of the system.

WP9 has as specific objectives:

- To support partners in identifying and protecting the IP of the project results.
- To develop an exploitation plan for the main project outputs, promoting and accelerating their uptake
- To ensure sustainability of the data exploitation infrastructure.

D9.3´s overall objective is to provide a structured interim strategy for exploitation and sustainability, as starting position acknowledging that some results may evolve in concept or relevance as they mature.

2.3 Methodology

The methodological approach includes the following steps:

1. Alignment with project framework: The report builds on the expected KERs described in the DoA and considers the commitments defined in the Grant Agreement regarding access, reuse and sustainability of clinical data and models.
2. Use of stakeholder insights: The report incorporates findings from D9.2 (interviews and questionnaires with internal and external stakeholders) regarding ARTEMIS' value proposition. In addition, information gathered through the KER questionnaire completed by partners during the Heidelberg Consortium Meeting (April 2025) was integrated.

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. Critical assessment of expected KERs: The expected KERs are reviewed to assess their maturity at M24, degree of innovation, exploitation potential and expected impact.
4. Prioritisation of KERs for detailed analysis: Based on this assessment, KER1, KER2 and KER3 are selected for deeper examination due to their higher development maturity and their relevant potential impact.
5. Mapping of underlying results: For each prioritised KER, we have identified the group of individual results that constitute it. This provides the basis for later decisions on IPR ownership and protection.
6. Functional, technological and market analysis: The selected KERs were examined to describe their role in supporting modelling and clinical use cases, how they compare with the current state of the art and their potential markets (including the identification of main players, main products and main barriers)
7. Outlining the principles for sustainability of data and models access for secondary use: work was conducted to develop a sustainability strategy for the ARTEMIS cohort beyond the duration of this project within the context of relevant initiatives currently under development such as European Health Data Space or the Advanced Platform of the Virtual Human Twin initiative.

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 Critical Assessment of Key Exploitable Results (KERs)

A KER is an identified main interesting result, which has been selected and prioritized due to its high potential to be "exploited", meaning to make use and derive benefits- downstream the value chain of a product, process or solution, or act as an important input to policy, further research or education.

3.1 Initial identification and description of key exploitable results

The expected KERs from ARTEMIS project are listed in the table below, as identified in the DOA and in D9.1.

Table 1. KER expected from ARTEMIS Project

Result ID	Title	Planned delivery date	Owners	Description
KER1	Processing pipelines for omics and digital imaging data, for the extraction of quantitative metrics as inputs to the multiscale models	**Q4 2025 – Q2 2026**	HULAFE IMPERIAL VHIR, DKFZ, ULEI	Harmonised and validated processing pipelines for imaging, digital pathology and -omics datasets, enabling extraction of standardised quantitative metrics to feed mechanistic and AI models.
KER2	Multi-modal clustering of patients and pattern identification	**Q3 2026**	BU	Algorithms for discovering clinically meaningful patient subgroups using multimodal datasets (clinical, imaging, -omics), supporting model development and interpretation.
KER3	Multiorgan multiscale multilevel model of the liver–heart axis for MAFLD (Virtual Twin models)	**Q4 2026** – Q2 2027	INRIA, DKFZ, ALU-FR, ULEI, MUV	Integrated mechanistic and AI-enabled Virtual Twin models representing liver–heart pathophysiology to support personalised predictions of disease progression and treatment outcomes.
KER4	MAFLD patient multimodal data visualisation module of ARTEMIs CDSS	Q1 2027	MEDEX	Visual analytics module enabling clinicians to explore multimodal patient data and model outputs within the ARTEMIs CDSS.
KER5	ARTEMIs CDSS	Q3–Q4 2027	MEDEX, BU, INRIA, DKFZ, ALU-FR, ULEI, MUV	Clinical decision support system integrating the Virtual Twin models, multimodal analytics, and patient risk stratification tools for MASLD/MASH management.
KER6	Federated data exploitation infrastructure with variables processed to	**Q4 2025 – Q3 2026**	MEDEX, APHP, JUH, VHIR, UKHD, CUSL, HULAFE, CHA,	Hybrid central/federated infrastructure for secure processing, mapping to the common data model (OMOP), and federated learning across all participating clinical sites.

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 common data model		ULS, MUV, SHEBA, BU	
KER7	Methodologies for assessment of trustworthiness (accuracy, robustness, reproducibility, explainability) in model performance	Q3 2027	INRIA, ALU-FR, ULEI, MUV, BU	Framework for evaluating and documenting the trustworthiness of mechanistic, AI and hybrid models, including robustness testing, reproducibility metrics and explainability tools.

The actual progress in achieving the expected results is monitored throughout the project. Any deviations in the project implementation (e.g. cases where certain KERs are only partially achieved, or results initially considered less relevant that eventually are of high value) will be reflected in the following versions of the KER list. These updates will be included in the periodic reports and the final exploitation plan.

3.2 Prioritization of key exploitable results for detailed assessment

The prioritization of KERs for detailed assessment analyses has been done based on their development status in M24, as well as their degree of Innovation, exploitability impact. These aspects are summarised as ·Attributes" of each KER in the table below.

Table 2. KERs development maturity and degree of innovation, exploitability and impact

KER 1	**Processing pipelines for omics and digital imaging data**
Type	Technical / Scientific result (software pipelines + methodology)
Description	The processing pipelines enable quantitative feature extraction in radiological imaging, digital pathology and omics data. They extract structured quantitative variables required for parameterisation, calibration and validation of the ARTEMIs models. Includes segmentation, feature extraction, radiomics/pathomics, omics quantification.
Value proposition	Enables reproducible transformation of heterogeneous clinical datasets into model-ready quantitative features, reduces site-to-site variability, accelerates model development, supports large-scale multicentre analysis and Virtual Twin personalisation.
Target audience	Clinical and biomedical researchers, modellers in academia and industry.
Planned delivery date	Q4 2025 – Q2 2026
Attribute	
Development status in M24	Medium, the core pipelines are under development, with initial segmentation and extraction workflows in place.
Degree of Innovation	High, the developments include integration of DL/ML segmentation and radiomics/pathomics and omics workflows. They address scanner/site heterogeneity.
Exploitability	High, they could be exploited as standalone pipelines, integrated in modelling workflows, or provided as a clinical/research processing service.
Impact	Medium, related to making more robust data available for modellers and researchers, thus enabling more robust research results and more accurate models.
KER 2	**Multi-modal clustering of patients and pattern identification**
Type	Algorithmic / Scientific result (AI/ML methods)

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.

Description	A set of clustering and pattern-detection algorithms that integrate clinical, imaging and omics variables to identify biologically and clinically meaningful MASLD/MASH patient subgroups. Supports discovery of phenotypes, risk strata and mechanistic hypotheses, and informs model calibration.
Value proposition	Reveals hidden structure in complex multimodal datasets to facilitate targeted modelling strategies and/or support personalised risk profiling.
Target audience	Clinical and biomedical researchers, modellers in academia and industry.
Planned delivery date	Q3 2026
Attribute	
Development status in M24	Low, corresponding to an early development using open data, with some initial clustering approaches tested.
Degree of Innovation	Medium, uses existing clustering approaches to stratify the patients in ARTEMIS cohort.
Exploitability	Medium, applicable to model calibration, CDSS stratification modules, biomarker discovery and cohort selection in clinical studies.
Impact	Medium, it may enabl new insights into MASLD phenotypes.
KER 3	**Multiorgan multiscale multilevel VT models of the liver–heart axis**
Type	Multiorgan, multiscale, multilevel Virtual Twin integrating liver–heart interactions
Description	Complex scientific/technical result (mechanistic + AI hybrid Virtual Twin system)
Value proposition	A suite of interconnected mechanistic and AI-supported models describing physiological, metabolic, haemodynamic and structural interactions between liver, heart and systemic circulation. Enables personalised prediction of disease progression, cardiovascular complications and simulated treatment effects.
Target audience	Provides a virtual representation of individual patients, offers mechanistic insight into liver–heart crosstalk, enables in silico scenario testing, forms the scientific core of the ARTEMIs CDSS.
Planned delivery date	Q4 2026 – Q2 2027
Attribute	
Development status in M24	Low, with submodels under active development for each of the four Use Cases..
Degree of Innovation	High, they will combine multiple scales and multi-organ interactions. They will represent some of the first VT models of the liver–heart axis.
Exploitability	High in the research context, opening exciting avenues, with still a long pathway of clinical evaluation and regulatory certification before being commercially exploitable.
Impact	High, facilitating VT-assisted biomedical research in metabolic diseases, and, in a longer future, personalised disease management.
KER4	**MAFLD patient multimodal data visualisation module of ARTEMIs CDSS**
Type	Technical / Software module
Description	An interactive visualisation module within the ARTEMIs CDSS, allowing clinicians and researchers to explore patient multimodal data, including clinical, imaging, and omics variables.
Value proposition	Enhances interpretability of complex multimodal datasets

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.

Target audience	Clinicians, biomedical researchers, IT developers
Planned delivery date	Q1 2027
Attribute	
Development status in M24	Very low, in the design phase
Degree of Innovation	Medium, the module will integrate multiple modalities into a single interactive interface designed for clinical use
Exploitability	Medium, it will be deployed within ARTEMIs CDSS prototype, with adaptability for other settings
Impact	Medium, it will support clinician understanding of complex patient data
KER5	**ARTEMIs CDSS**
Type	Technical / Software module
Description	The ARTEMIs CDSS will combine mechanistic and AI-supported models and multimodal patient data for the 4 targeted use cases, to provide personalised disease predictions, risk assessment, and treatment simulations for these types of patient.
Value proposition	A platform for personalised medicine in the field of MASLD patient management
Target audience	Clinicians, biomedical researchers, IT developers
Planned delivery date	Q3–Q4 2027
Attribute	
Development status in M24	Very low, in the design phase
Degree of Innovation	High, it will integrate novel models based on the liver-heart axis VT.
Exploitability	Medium, it will be functional for the management of patients corresponding to the use cases, but also designed for future scalability and use in clinical settings
Impact	High, it will enable personalised treatment planning.
KER 6	**Federated data exploitation infrastructure with variables processed to the common data model**
Type	Technical / Infrastructure result
Description	A hybrid (federated + central) data infrastructure enabling secure exploration, analysis and processing of clinical, imaging, omics and digital pathology datasets across ARTEMIs sites, allowing cross-site federated learning and controlled access for modelling teams.
Value proposition	Provides ARTEMIs partners with a workable and compliant environment to harmonise and access multimodal datasets without full centralisation. Facilitates internal project workflows by enabling consistent data structures for modelling teams.
Target audience	Clinical partners, modellers, AI developers, EU research infrastructures, hospitals implementing federated analytics,.
Planned delivery date	Q4 2025 – Q3 2026
Attribute	

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.

Development status in M24	Medium, a central cloud-based in deployed and the federation of other nodes is in progress.
Degree of Innovation	Low , the infrastructure reuses established standards (OMOP) and existing architectural patterns already implemented in other data sharing infrastructures.
Exploitability	Low, the platform is tailored to the ARTEMIs consortium's needs with limited transfer potential outside the project.
Impact	Low , the infrastructure supports internal project needs, its contribution is facilitating the use of the data by the modellers, using state of the art solutions.
KER 7	**Methodologies for assessment of trustworthiness in model performance**
Type	Methodological / Scientific result
Description	A set of evaluation frameworks and metrics designed to assess the reliability, robustness, and generalisability of ARTEMIs predictive models and Virtual Twin outputs. Includes bias assessment, uncertainty quantification, and validation protocols across multimodal datasets.
Value proposition	Ensures confidence in model predictions
Target audience	Model developers
Planned delivery date	Q3 2027
Attribute	
Development status in M24	Very low
Degree of Innovation	Medium, it will integrate established validation principles with ARTEMIs-specific modelling challenges
Exploitability	Low-Medium, it will be applicable for internal model validation and potentially for certification.
Impact	High, it will increase credibility and adoption potential of ARTEMIs predictive models

As summarised in Table 3, the development status, degree of innovation, exploitability and impact of each KER indicate that KER1, KER2 and KER3 are the most mature and actionable results, with relevant innovation and societal impact potential. For this reason, sections 4, 5 and 6 of this report are devoted to the assessment of KER1, KER2 and KER3, respectively.

Table 3. Prioritization of KERs for detailed assessment

KER	Development status in M24	Degree of		
		Innovation	Exploit.	Impact
1 Processing pipelines	Medium	High	High	Medium
2 Patient clustering	Low	Medium	Medium	Medium
3 Models the liver–heart axis	Low	High	High	High
4 Visualisation module	Very low	Medium	Medium	Medium
5 ARTEMIS CDSS	Very low	High	Medium	High
6 Federated infrastructure	Medium	Low	Low	Low
7 Methodologies for assessment of trustworthiness	Very low	Medium	Low	High

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 KER1 – Processing pipelines for extraction of quantitative metrics as inputs to the models

4.1 Functional analysis

The development of *processing pipelines for extraction of quantitative metrics as inputs to the models* responds to broad scientific, clinical and market needs that are not limited to ARTEMIS. They are central to the advancement of personalised medicine, large-scale clinical research and computational biology, where virtual twins, mechanistic models and AI systems increasingly depend on high-quality, model-ready patient data.

Modern healthcare routinely generates heterogeneous raw data: digital images (e.g. MRI, CT), electronic health records, and high-throughput omics outputs (proteomics, metabolomics, lipidomics). On their own, these data streams are not directly usable by computational models. There is a growing need for robust pipelines that systematically transform this complexity into standardised, quantitative features.

4.1.1 Extraction of quantitative data for model development

The first need is to convert heterogeneous clinical and biological signals into structured, quantitative inputs that can drive mechanistic and AI models. These models underpin virtual twins and other advanced decision support tools. Processing pipelines bridge the gap between raw formats (e.g. imaging series, mass spectrometry outputs, sequencing data) and the numerical variables required by models at different scales and abstraction levels. This includes:

- Quantitative feature extraction from omics and digital pathology (e.g. protein or metabolite abundances, tissue composition metrics), providing model parameters and validation targets.
- Segmentation and quantification for simulation, by deriving geometric and topological descriptors of organs, tissues and lesions that are required to initialise and personalise spatially and temporally resolved simulations.

Without this intermediate layer, model developers must rely on ad hoc scripts or small, curated datasets, which limits scalability, reproducibility and clinical translation.

4.1.2 Standardisation and harmonisation of multimodal, multicentre data

A second need arises from the way clinical research is organised. Large, collaborative, multicentre studies are essential to achieve statistical power, cover the full disease spectrum and ensure representativeness of real-world populations. However, data acquired at different sites and with different technologies suffer from technological and methodological heterogeneity:

- Different scanners, imaging protocols and laboratory workflows introduce systematic differences that bias downstream analyses.
- Data are stored in heterogeneous systems (EHRs, PACS, laboratory information systems) and described using different terminologies.

Processing pipelines are needed to implement robust, standardised pre-processing and harmonisation steps (such as bias field correction, resampling, intensity normalisation and denoising in imaging, or controlled vocabularies and quality thresholds in omics) and to make

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.

datasets interoperable with common data models and standards. This is a prerequisite for trustworthy model development, validation and reuse across centres and vendors.

4.1.3 Extraction of translational biomarkers and radiomic signatures

A third driver is the demand for clinically meaningful, quantitative biomarkers that can be derived from routinely collected data. Health systems and industry are increasingly interested in imaging-based and omics-based biomarkers that can support early diagnosis and risk stratification, longitudinal monitoring of disease progression and treatment response, or enrichment and endpoint definition in clinical trials.

Processing pipelines enable the systematic extraction of such biomarkers, for instance by:

- Deriving imaging metrics such as proton density fat fraction, or elastography-based stiffness, and high-dimensional radiomics features from relevant imaging sequences.
- Quantifying tissue architecture, cell densities and molecular marker expression in digital pathology, replacing or complementing subjective and time-consuming manual assessments.

These outputs are valuable both as direct clinical indicators and as inputs or validation targets for models.

4.1.4 Supporting data-driven discovery and clinical decision support

Finally, reliable, standardised quantitative outputs from processing pipelines are essential to unlock the value of large cohorts for research and clinical decision support:

- For data-driven discovery, they support robust exploratory analyses, clustering and pattern identification across large populations, which is critical for hypothesis generation, endotype discovery and refinement of mechanistic understanding.
- For clinical decision support, they provide the validated, traceable inputs that computational models and AI-based tools require to produce clinically meaningful predictions and recommendations.

Overall, there is a clear and growing scientific and market demand for reusable, standardised processing pipelines that can convert heterogeneous patient data into quantitative, interoperable metrics suitable for multiscale models and clinical applications, across diseases and care settings.

4.2 Technology Analysis

4.2.1 Limitations of the current state of the art

While recent years have seen significant progress in quantitative imaging (e.g., IBSI[1]), clinical data standardization (e.g., OMOP CDM[2]), and the rise of commercial multimodal platforms (e.g., SOPHiA DDM, Siemens AI-Rad Companion), the current landscape still presents critical limitations.

Despite this rapid expansion in concepts and prototypes, these advances have not yet solved the challenge of deploying systematic, interoperable processing pipelines. Key gaps remain in providing robust, standardized quantitative inputs for multiscale mechanistic and AI models (like

[1] https://theibsi.github.io

[2] https://www.ohdsi.org/data-standardization/

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.

virtual twins) in federated, multicentre settings. These limitations can be grouped into four main areas:

4.2.1.1 Residual fragmentation and partial standardisation of multimodal data

On the clinical data side, the OMOP CDM and its ecosystem enable the transformation of disparate observational datasets into a common structure and representation, including standardised vocabularies and analytic routines[3]. However, implementing these standards requires complex Extract-Transform-Load (ETL) pipelines, local terminology mapping and rigorous data quality assessment. In many hospitals and research networks, these processes remain incomplete or cover only part of the available information (e.g. structured *Electronic Health Record* - EHR - fields but not imaging metadata, unstructured text or high-dimensional omics).

In parallel, commercial multimodal platforms such as SOPHiA DDM™ for Multimodal[4] show that it is technically feasible to ingest, harmonise and analyse genomic, imaging and clinical data within a unified environment for precision oncology and biopharma collaborations. However, these platforms are proprietary, indication-focused and optimised around their own internal data models. Their processing pipelines are not openly reusable, and they are not designed as generic, transparent input layers for academic multiscale models or virtual twin frameworks across diseases.

Consequently, real-world clinical and biological data remain fragmented across EHR, PACS, laboratory and omics systems, with only partial standardisation and limited support for the systematic extraction of model-ready variables that can be shared and reused across institutions.

4.2.1.2 Persistent challenges in cross-site robustness and reproducibility of quantitative imaging pipelines

In quantitative imaging, IBSI has provided a robust foundation by standardising feature definitions and validating reference values for radiomics features across CT, PET and MRI, which improves comparability and calibration of different software implementations[5]. Nevertheless, multicentre studies consistently show that scanner-, proto col- and site-related variability remains a major source of bias in radiomics and quantitative image analysis. A recent review of harmonisation strategies in MRI-based radiomics concluded that, although numerous image- and feature-level harmonisation methods exist (including ComBat and its variants), none has yet emerged as a universally accepted solution, and residual differences between centres are still common[6].

Clinical AI products in radiology, such as AI-Rad Companion[7], provide automated post-processing and organ-specific quantification for a range of modalities and indications, integrated in the radiologist's workflow and cleared as medical devices. However, these tools are typically focused on narrow tasks (e.g. organ volumetry or detection of specific findings), are closely tied to particular vendors and modalities, and expose only high-level results. They do not generally offer open, configurable feature-extraction pipelines or detailed quantitative outputs that can be

[3] https://ohdsi.github.io/CommonDataModel/

[4] https://www.sophiagenetics.com/sophia-ddm/

[5] Zwanenburg, A., Vallières, M., Abdalah, M. A., Aerts, H. J., Andrearczyk, V., Apte, A., ... & Löck, S. (2020). The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology, 295(2), 328-338. https://doi.org/10.1148/radiol.2020191145

[6] https://www.mdpi.com/2313-433X/8/11/303

[7] https://www.siemens-healthineers.com/digital-health-solutions/ai-rad-companion

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.

systematically reused as inputs or validation targets for multiscale, research-grade models across centres.

Thus, despite substantial progress, there is still no widely adopted, open framework that guarantees robust and reproducible quantitative imaging pipelines across scanners, vendors and use cases capable of feeding multiscale modelling efforts in a standardised manner.

4.2.1.3 Limited availability of truly multimodal, longitudinal datasets ready for multiscale modelling

Meta-reviews and scoping reviews on digital twins in healthcare highlight a rapidly growing but heterogeneous body of work, with many examples of organ-specific models and conceptual architectures, yet relatively few large-scale, clinically integrated implementations[8]. Most existing digital twins or advanced mechanistic models: 1) are centred on a single organ or disease area, 2) rely on small or monocentric cohorts, or on synthetic data, and 3) use only a subset of potentially available modalities (e.g. clinical parameters plus imaging, or clinical parameters plus biomechanics), with incomplete integration of omics, imaging, treatment histories and outcomes over time.

Platforms such as SOPHiA DDM™ for Multimodal demonstrate the construction of large multimodal datasets that unify genomic, imaging and clinical oncology data for advanced analytics. Yet these datasets are typically not accessible as general-purpose infrastructure for external mechanistic or multiscale modelling and are not natively aligned with open common data models or explicit "model input" specifications.

Consequently, multiscale models and virtual twins that aim to integrate anatomy, physiology, molecular profiles and longitudinal clinical trajectories still lack access to harmonised, multimodal datasets of sufficient size, diversity and standardisation. The bottleneck is not only data availability, but also the absence of open, well-documented pipelines that systematically transform raw multimodal inputs into a consistent, model-aligned set of quantitative metrics.

4.2.1.4 Weak integration of processing pipelines with modelling workflows and clinical systems

There is also a structural gap between existing data processing tools and the end-to-end workflows required for virtual twins and advanced decision support systems. A recent meta-review on digital twins in healthcare concludes that, although conceptual architectures and use cases are proliferating, only a minority of implementations are tightly integrated with hospital information systems, real-time data streams and clinical decision processes[9].

On the operational side, many radiomics and omics pipelines remain research-grade scripts or bespoke workflows tied to specific projects or centres, with limited modularity, version control and documentation. Commercial platforms often provide dashboards and high-level analytic outputs, but do not expose low-level, configurable processing modules and interfaces that would allow seamless integration with external mechanistic or AI models, or with federated learning and privacy-preserving analytics frameworks.

---

[8] https://www.jmir.org/2025/1/e69544

[9] Ringeval, M., Etindele Sosso, F. A., Cousineau, M., & Paré, G. (2025). Advancing health care with digital twins: meta-review of applications and implementation challenges. Journal of Medical Internet Research, 27, e69544. https://doi.org/10.2196/69544

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.

From a clinical integration perspective, deployment of AI and quantitative tools in radiology and pathology is still uneven, and these tools are rarely conceived as systematic upstream providers of standardised model inputs for multiscale simulations or virtual twins.

4.2.1.5 Analysis and conclusions of the state of the art

Taken together, the current state of the art shows clear progress in standardisation, harmonisation and multimodal analytics, but also reveals key gaps for ARTEMIS: there is a lack of interoperable, open and reusable processing pipelines that 1) operate robustly across centres and technologies, 2) transform heterogeneous multimodal patient data into a consistent set of quantitative metrics, and 3) integrate natively with multiscale mechanistic and AI models in virtual twin applications. These gaps motivate the development of **KER1 – Processing pipelines for extraction of quantitative metrics as inputs to the models** within ARTEMIS.

4.2.2 ARTEMIS results and technical features beyond the state of the art

Work Package 3 (WP3) implements the "Processing pipelines for extraction of quantitative metrics as inputs to the models" (KER1) by transforming raw digital pathology, imaging and multi-omics data into harmonised, quantitative inputs for the multiscale mechanistic and AI models.

The table below consolidates the main WP3 results contributing to KER1, grouped into 11 components (R1–R11). For each result, it highlights the core functionality, the technical features that go beyond the state of the art, the main contributing partners and the potential for further research or commercial exploitation.

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 4. Prioritization.Main results clustered under KER1

ID	WP3 Task(s)	Main partners contributing	Result / component	Description and functionality	Technical features beyond the state of the art	Exploitation potential and link to KER1
R1	T3.1, T3.4	ULEI (lead), HULAFE, INRIA, UKHD, AP-HP, MEDEX	**Digital pathology and liver microarchitecture quantification pipeline (TiQuant)**	Refined processing and quantification pipelines for 2D/3D liver histology, based on confocal imaging and multiple stains (CD45, PERLS, Sirius Red, Adipophilin), delivering architectural and dynamic parameters (vessel branching, sinusoid and bile canaliculi diameters, cell densities, biomarker scores).	Builds on the established TiQuant framework[10] but extends it with multi-marker, quantitative digital pathology specifically tuned to chronic liver disease and MASLD, enabling systematic extraction of microarchitectural metrics at scale rather than manual grading.	Can be packaged as an advanced digital pathology analytics module for research hospitals and pharma/CROs. Within KER1, it provides histology-derived quantitative parameters for microscale liver models (T5.4) and a reference for mapping non-invasive biomarkers to tissue-level changes.
R2	T3.1	HULAFE (lead), ULEI	**CT segmentation suite and centralised model zoo – CT branch**	Integrated CT segmentation pipeline built as the CT branch of the ARTEMIS model zoo. It combines TotalSegmentator[11] with pre-trained liver and HCC models, packaged in Docker for straightforward deployment	Provides a curated model zoo of nnU-Net–based[12] CT models with unified configuration and versioning, enhancing consistency for multicentre/multiscanner data. ULEI provides the	Delivers standardised organ and lesion masks for volumetry, radiomics, and geometric inputs to liver/flow Virtual Twin models. Exploitable as a cloud service or in imaging core labs.

[10] https://academic.oup.com/bioinformatics/article/31/19/3234/211701

[11] TotalSegmentator is a free open-source tool developed by the Research dept. at the University Hospital of Basel https://totalsegmentator.com

[12] https://www.nature.com/articles/s41592-020-01008-z

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 validated by expert radiologists.	centralized Model Zoo infrastructure (R3, T3.1)	
R3	T3.1	HULAFE (lead), ULEI, MUV	**MRI segmentation suite and centralised model zoo – MRI branch**	Suite of MRI segmentation models for liver, HCC, spleen, kidney, and visceral / subcutaneous fat. Models cover multi-echo, mDIXON, and T1/T2 sequences, forming the MRI branch of the ARTEMIS model zoo.	Delivers a multi-organ, multisequence nnU-Net–based model zoo for abdominal MRI, providing an extensible framework that moves beyond isolated, single-task models	Forms the core of an MRI liver and body composition quantification package, supplying standardised MRI masks feeding mechanistic VT submodels and ML components
R4	T3.2	HULAFE (lead), ULEI, MUV	**Fat fraction quantification pipelines (MRI PDFF/FF and CT HU)**	Standardised procedures to obtain Proton Density Fat Fraction (PDFF) or Fat Fraction (FF) from MRI under three acquisition scenarios (multi-echo magnitude/phase, In/Out-of-phase or Water/Fat images, scanner-generated PDFF/FF maps) and to estimate fat content from CT using Hounsfield Unit thresholding.	Translates scattered methods into robust, containerised pipelines that can adapt to heterogeneous MRI protocols and CT acquisition, while enforcing consistent derivation of PDFF/FF and HU-based fat metrics across centres. This directly addresses SOTA variability in fat quantification workflows.	Immediately exploitable as clinical or research pipelines for liver steatosis and body composition quantification, integrable into PACS/post-processing systems. For KER1, they deliver core imaging biomarkers (steatosis, myosteatosis) used for model parameterisation and validation.
R5	T3.2	HULAFE (lead), ULEI, MUV, MEDEX	**Radiomics and deep-feature pipelines (CT/MRI, including HCC)**	Docker-based pipelines for radiomics and deep-feature extraction from selected CT and MRI series, including pre-processing, feature computation and preparation	Combines radiomics features with deep features from CNNs and RadImageNetwithin reproducible, containerised workflows, and integrates multi-parametric MRI (DCE,	High potential as a generic radiomics service for multicentre trials and registries. Within KER1, these pipelines provide standard feature-extraction blocks for

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.

				of feature sets for clustering and pattern recognition (e.g. for HCC multi-parametric MRI datasets).	PDFF, DWI/ADC, T2, MRE) for HCC, going beyond single-sequence radiomics studies.	WP4 clustering and WP5 ML submodels (e.g. disease course prediction and risk stratification in liver disease).
R6	T3.3	HULAFE (lead), MUV, MEDEX	**CT/MRI preprocessing and harmonisation pipeline**	Implemented preprocessing pipelines to reduce scanner- and protocol-induced variability in CT and MRI, including organ-centred HU clipping, denoising, contrast enhancement, bias field correction, spatial rescaling and intensity normalisation across vendors and protocols.	Delivers a systematic, liver-focused harmonisation pipeline for CT/MRI, explicitly designed for multicentre/multivendor data in line with the scanner heterogeneity problem identified in the SOTA section, rather than ad hoc pre-processing per study.	Core cross-cutting KER1 module: a generic pre-processing step reusable across all imaging pipelines and future projects. It can evolve into a harmonisation plugin for radiology departments and imaging core labs that need robust, standardised quantitative imaging inputs.
R7	T3.3	MUV (lead), HULAFE	**Domain adaptation and continual learning framework for MRI analysis**	Framework to adapt and maintain image analysis algorithms across sites and changing scanner characteristics, using domain adaptation and rehearsal-based continual learning strategies for liver MRI analysis.	Goes beyond static domain adaptation by addressing temporal drift and evolving imaging devices, leveraging rehearsal methods to select training data that maximise diversity of imaging characteristics and investigating how disease properties are encoded in latent representations.	Offers a generic methodology and codebase for vendors and research networks to keep AI models valid over time and across centres. For KER1, it provides a path to sustain the performance of segmentation and feature-extraction pipelines as the ARTEMIS imaging ecosystem evolves.

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.

R8	T3.4	ULEI (lead), HULAFE, UKHD, AP-HP, INRIA, MEDEX	**Digital histology biomarker panel and microstructure metrics**	Set of histological biomarkers and quantitative metrics for pathological substrates (CD45, PERLS, Sirius Red, Adipophilin), normalised to reference samples and linked to specific microstructural changes (e.g. microsteatosis vs vacuoles, iron deposition patterns).	Moves from qualitative scoring to standardised, quantitative biomarker panels that can be systematically derived from digital slides, providing a richer parameterisation of liver microenvironment and fibrosis progression than traditional semi-quantitative grades.	Valuable as a digital pathology biomarker package for MASLD/CLD studies and pharma trials. In KER1, these metrics act as ground truth for microscale models and as targets for surrogate relationships with non-invasive imaging and blood biomarkers.
R9	T3.4	INRIA (lead), ULEI	**Generative and reconstruction tools for liver microarchitecture**	Concept and initial tools to generate artificial liver lobule architectures calibrated to MASLD progression, conceptual strategy to infer histological parameters from non-invasive modalities, and first promising 3D microarchitectural staining for human liver.	Addresses a key gap in SOTA by proposing a parametric generator of liver microarchitecture and a conceptual mapping from imaging/biomarkers to histology-level parameters, enabling synthetic microstructures and "out-of-image" disease progression modelling.	High potential as a research tool for multiscale modelling groups and future IPR. Within KER1, it defines a route to generate model-compatible microarchitectures when histology is sparse, and to explore non-invasive surrogates of microstructural change.
R10	T3.5	DKFZ (lead), VHIR	**Automated quantitative proteomics pipeline and liver proteome dataset**	Automated SP3-based sample preparation pipeline on a Bravo liquid handling robot for tissue and liquid biopsies, combined with DIA global proteomics and PRM-based targeted	Integrates high-throughput, automated sample preparation with a standardised DIA/PRM workflow tailored to ARTEMIS, providing	The workflow and dataset can be exploited as a reference proteomics pipeline and cohort for future liver disease studies and VT projects. For KER1, it supplies molecular-

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.

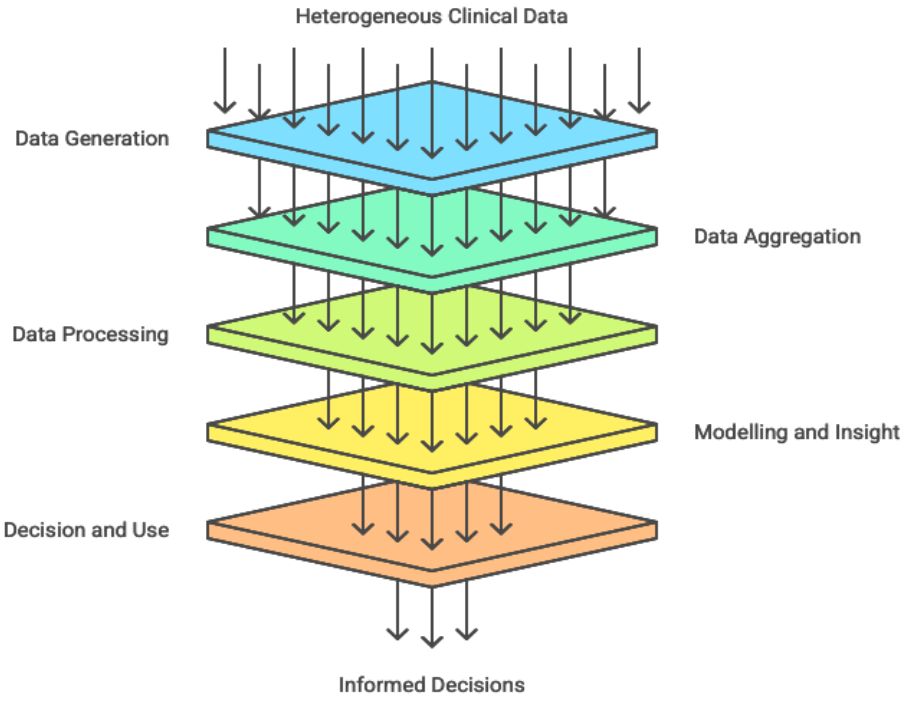
				quantification applied to 62 liver tissue samples (7,932 proteins identified).	quantitative protein abundance data directly aligned with VT mechanistic modelling requirements rather than generic biomarker discovery.	level metrics for intracellular signalling and pathway models (T5.3) and for multimodal data fusion (WP4).
R11	T3.6	VHIR (lead), DKFZ	**Standardised metabolomics and lipidomics pipeline (serum and tissue)**	Procedures and SOPs to handle ARTEMIS samples using an existing NMR-based metabolomics and lipidomics pipeline (Bruker Avance II 600 MHz, CPMG sequence), with internal calibration, analytical thresholds and protocols for polar and non-polar metabolite extraction from serum and liver tissue.	Formalises a project-wide, quality-controlled pipeline for metabolomics and lipidomics, including logistics (shipment, storage) and analysis steps, ensuring consistent quantitative profiles across sample types and time points, beyond typical single-study implementations.	Can be reused as a standard metabolomics / lipidomics operating package for future cohorts in liver and cardiovascular research. Within KER1, it delivers quantitative metabolite and lipid metrics that complement imaging, histology and proteomics in the multimodal VT parameterisation chain.

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.

4.3 Market analysis

4.3.1 Value chain overview

The value chain surrounding KER1's quantitative processing pipelines begins with Data Generation at clinical sites, including hospitals and laboratories. This process creates a vast array of raw, heterogeneous, and siloed data, such as DICOM images. This data then flows into Data Aggregation & Storage systems, which are typically vendor-locked and fragmented (e.g., PACS). KER1's primary role is at the next step: Data Processing & Quantification. It acts as a harmonisation and quantification engine, pulling data from these upstream silos, applying its standardised pipelines, and transforming the raw data into interoperable, model-ready quantitative metrics. These metrics are the value-added product that flows to the Modelling & Insight Generation layer, where the metrics are used. The final step is Decision & Use, where the model outputs are translated into actionable insights, such as R&D decisions for pharma, new hypotheses for researchers, or, in the future, clinical decisions for physicians.



The first link in the value chain is **Data Generation**. The primary actors at this stage are the clinical and research sites, including university hospitals, liver centers, and specialized laboratories. Their core activity involves the initial creation of patient data, encompassing imaging studies (CT, MRI), digital pathology scans, omics assays, and EHR records. This process results in an output of technically rich, diverse raw data. However, this output is characterized by high fragmentation and significant variability due to disparate protocols, scanners, and workflows.

The second link in the value chain is **Data Aggregation**. The primary actors at this stage are hospital IT departments, research infrastructures, and technology vendors (e.g., PACS, EHR, LIS providers). Their core activity involves collecting and storing the raw data from the generation sites using systems like Vendor Neutral Archives, EHR platforms, and research repositories. This process results in an output of aggregated, centralized data inventories. However, this output is characterized by data silos, heterogeneous models, and vendor lock-in, which hinders downstream accessibility and integration.

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 third link in the value chain is **Processing & Quantification**. This is the layer where KER1 creates most of its value. Its core activity involves pulling the fragmented data from the aggregation layer and applying its suite of standardized pipelines for imaging, digital pathology, and omics (e.g., harmonization, segmentation, feature extraction). This process results in an output of robust, standardized, and model-ready quantitative metrics. This output is characterized by being coherent, traceable, and explicitly aligned with the input requirements of advanced analytical models.

The fourth link in the value chain is **Modelling & Insight Generation**. The primary actors at this stage are mechanistic modelers, AI/ML teams, biostatisticians, and digital twin developers. Their core activity involves consuming the harmonized metrics from KER1 to build, calibrate, train, and validate sophisticated analytical tools. This process results in an output of virtual human twins, predictive AI models, and novel prognostic signatures. This output is characterized by enhanced reproducibility and reliability, as the standardized inputs from KER1 reduce systemic bias and variability.

The final link in the value chain is **Decision & Use**. The primary actors at this stage are the end-users, including pharma R&D teams, CROs, clinicians (hepatologists, radiologists), regulators, and payers. Their core activity involves applying the generated models and insights to high-stakes challenges, such as clinical trial enrichment, patient stratification, and advanced diagnostic support. This process results in an output of data-driven clinical and research decisions. This output is characterized by its real-world impact and provides a critical feedback loop for refining the upstream KER1 processing pipelines.

4.3.2 Potential application fields for the KER

Given ARTEMIS' focus on liver disease (MASLD/MASH and related cardio-metabolic conditions), the most relevant application fields are those where quantitative liver imaging, digital pathology and omics-derived biomarkers can materially reduce trial costs, improve endpoint robustness, or enable multi-scale modelling. Below, we summarise these fields and then narrow down to the segments with the highest near-term commercial potential in Europe.

4.3.2.1 Pharma / CRO clinical trials in MASLD/MASH and related liver disease

The most attractive application field for KER1 is quantitative biomarker generation for pharma and CRO-run clinical trials in MASLD/MASH and other chronic liver diseases. In these trials, liver biopsies remain the gold standard endpoint but are invasive, costly and subject to high inter-reader variability. Regulators and sponsors are therefore actively exploring quantitative imaging biomarkers (QIBs) and AI-supported histology assessment to de-risk programmes and speed recruitment. The recent EMA endorsement of AIM-NASH, an AI tool to assess MASH histology in trials, illustrates that regulators are willing to accept advanced image-based tools when variability is demonstrably reduced[13].

KER1's liver-focused MRI/CT pipelines, digital pathology metrics and omics workflows map directly onto three high-value services for this segment: 1) non-invasive liver fat/fibrosis quantification for enrichment and response monitoring, 2) digital pathology quantification of biopsies as a more reproducible endpoint, and 3) multimodal feature panels that support mechanistic understanding and patient stratification.

[13] https://www.reuters.com/business/healthcare-pharmaceuticals/eu-health-regulator-clears-use-ai-tool-fatty-liver-disease-trials-2025-03-20

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.

Demand is underpinned by the epidemiology of MASLD: recent analyses estimate that around 32–38 % of adults globally meet MASLD criteria, with prevalence particularly high among patients with type 2 diabetes[14]. The MASLD/MASH trial pipeline has expanded rapidly over the last decade, with over 1,200 interventional MASLD/MASH trials registered in the ClinicalTrials.gov database by 2024[15]. For KER1, the immediate opportunity is to offer "research-use-only" (RUO) imaging-pathology-omics analysis services to sponsors and CROs, without requiring full MDR certification by 2027 (assuming a TRL5 "demonstration in operating environment").

On the market side, the global quantitative imaging biomarker services market is estimated at USD 1.47 billion in 2024, with a projected CAGR of 13.7 % to 2033[16]. The global clinical trial imaging market adds a further USD 1.3–2.6 billion, growing at around 7–9 % annually[17]. Europe accounts for roughly EUR 0.4 billion of clinical trial imaging spend in 2025, with a 5.5 % CAGR[18]. Assuming Europe represents ~25–30 % of global QIB services, the addressable EU-27 market for KER1-type quantitative services is on the order of EUR 300–400 million per year.

4.3.2.2 Multicentre clinical research consortia and registries (liver and cardio-metabolic)

A second core field is large clinical research consortia and registries focused on liver and cardio-metabolic disease across Europe. These include EU-funded initiatives and hospital networks adopting common data models such as OMOP (e.g. OHDSI-Europe, EHDEN) to pool EHR and imaging data at scale. Their main bottleneck is not data storage but the lack of robust, reproducible pipelines to harmonise multicentre imaging (different scanners, vendors, protocols) and to link imaging, pathology and omics into a coherent research dataset.

KER1 directly addresses these pain points via its harmonisation and domain-adaptation pipelines for CT/MRI and its standardised extraction of liver-centric imaging and histology metrics. For principal investigators and data-intensive hepatology research units, KER1 can be deployed as a containerised processing layer in existing Trusted Research Environments, producing reproducible metrics across sites and over time, which is critical for longitudinal liver cohorts and cardio-hepatic interaction studies.

Financially, this field does not offer the same per-study budgets as pharma, but it represents a sustained software and services opportunity. The global clinical data management system (CDMS) market is projected at USD 3.46 billion in 2025 with an 11.1 % CAGR to 2034[19]. Europe accounts for roughly 25–30 % of CDMS spend (≈USD 0.9–1.0 billion). KER1 targets the higher-value niche of imaging/omics-intensive research data processing, plausibly 10–15 % of this European CDMS market, this yields an addressable EU-27 opportunity in the order of EUR 80–120 million per year, consistent with the internal market analysis.

4.3.2.3 Multimodal data and AI platforms (B2B channel)

A third field is not an end-customer segment but a strategic channel: commercial multimodal data and AI development platforms that provide compliant infrastructure and increasingly host

[14] https://pmc.ncbi.nlm.nih.gov/articles/PMC11925440/
[15] https://pmc.ncbi.nlm.nih.gov/articles/PMC11887356/
[16] https://dataintelo.com/report/quantitative-imaging-biomarker-services-market
[17] https://www.grandviewresearch.com/industry-analysis/clinical-trial-imaging-market
[18] https://straitsresearch.com/report/europe-clinical-trial-imaging-market
[19] https://www.precedenceresearch.com/clinical-data-management-system-market

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.

containerised analysis "apps". Examples include Flywheel (imaging-centric TRE), DNAnexus (clinico-omics TRE) and SOPHiA GENETICS' DDM platform. These players already serve pharma and research customers who face the same challenges ARTEMIS addresses: scalable, validated pipelines for quantitative liver metrics and multimodal fusion.

Here, KER1 would be packaged as an embedded processing engine (e.g., a set of Docker containers or APIs) that partners can offer as premium analysis modules in their PaaS/SaaS environments. The underlying market is the broader AI-in-medical-imaging and multimodal analytics segment, estimated at USD 1.36 billion globally in 2024 and projected to grow at around 35 % CAGR to 2033[20]. A realistic KER1 opportunity in Europe is a small fraction of this (tens of millions of euros annually in licensing or revenue-share across a handful of platforms), but it provides strong leverage: one technical integration can expose KER1 to many pharma and academic users.

4.3.2.4 Advanced clinical diagnostics in tertiary hepatology / cardio-hepatic centres

The final field of interest is the deployment of KER1-derived pipelines as regulated Software as a Medical Device (SaMD) in advanced clinical settings, especially tertiary centres with strong hepatology, radiology and cardiology services. Here, KER1-type tools would power quantitative liver reports (steatosis, fibrosis surrogates, body composition, liver and heart volumes) that feed into routine clinical decision making and multidisciplinary boards.

This is strategically attractive but constrained by regulatory, integration and workflow barriers. Competing AI-in-imaging SaMD solutions (e.g. organ volumetry and liver multiparametric MRI tools) are already entering the market, and the global AI in medical imaging segment is forecast to grow from USD 1.36 billion in 2024 to nearly USD 20 billion by 2033 (CAGR ~35 %)[21]. However, achieving MDR class IIb/III certification and deep integration with major PACS/EHR systems (Epic, Oracle Cerner, Sectra, GE, Philips) implies long timelines and high investment. Given KER1's expected TRL (demonstration in operating environment by 2027), clinical diagnostics should be treated as a second-wave opportunity, building on validation and revenue generated in research and trial settings.

4.3.2.5 Virtual Human Twin (VHT) and in-silico environments (emerging, research-driven)

While less immediately commercial, VHT and in-silico trial environments are strategically important for ARTEMIS. KER1 was explicitly designed to provide the multi-scale quantitative inputs (micro-architecture, organ-scale metrics, molecular profiles) required for virtual liver twins and cardio-hepatic models. The European Virtual Human Twin initiatives and related projects (e.g. ex-EDITH CSA) are consolidating funding streams for such infrastructures[22]. For KER1, this field is best viewed as a research exploitation path (participation in consortia, spin-off VHT projects) rather than a stand-alone commercial market in the short term.

[20] https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-medical-imaging-market
[21] https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-medical-imaging-market
[22] https://www.sciencedirect.com/science/article/pii/S166526812500122X

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.

4.3.3 Prioritisation and narrowing of target markets

Considering technical fit, unmet need, purchasing power and regulatory barriers, the segments with the highest potential for commercialisation of KER1 by 2027 are:

1. **Pharma / CRO clinical trials in MASLD/MASH and chronic liver disease (EU-focused)**: clear unmet need for non-invasive, reproducible liver biomarkers, strong willingness to pay, viable RUO/service model, and rapidly growing quantitative imaging and digital pathology spend.
2. **Multicentre clinical research consortia and registries in liver and cardio-metabolic disease**: strong technical alignment with ARTEMIS' origins, growing demand for harmonised imaging/omics pipelines within federated data networks, and recurrent, medium-ticket software/services revenues in Europe.
3. **Multimodal data & AI platforms as B2B channels**: not a primary source of demand, but a capital-efficient route to reach pharma and academic users, particularly for liver-focused quantitative processing "apps" integrated into existing TREs.

Advanced clinical diagnostics and VHT environments remain important *strategic* fields, but they are better positioned as second-wave opportunities after KER1 has achieved maturity and validation through the first two segments.

Table 5. Analysis of market segments for KER1

Segment	Main unmet need	KER1 offer (R1–R11)	Market characteristics	Priority (by 2027)
1. Pharma / CRO liver trials	Robust, non-invasive and histology-based endpoints, cross-site consistency	Liver QIBs (R2–R6), digital pathology (R1, R8, R9), multimodal feature panels (R5, R10, R11)	EU addressable market ≈ EUR 300–400 M/year, high growth	High – primary
2. Multicentre consortia & registries	Harmonised imaging/omics pipelines in federated datasets	Harmonisation (R6, R7), segmentation/radiomics (R2–R5), pathology/omics (R1, R8–R11)	EU opportunity ≈ EUR 80–120 M/year, recurrent project-based	High – strategic and recurrent
3. Multimodal AI platforms (B2B)	Liver-focused analytics "apps" in existing TREs	Embedded Docker/APIs for KER1 pipelines	Indirect exposure to a fast-growing global AI imaging market	Medium – leverage channel
4. Advanced clinical diagnostics	Quantitative liver reports integrated in routine care	Same core pipelines, but as SaMD modules	Attractive but regulatory-/integration-heavy	Low short-term – second wave
5. VHT / in-silico	Model-ready multimodal inputs for virtual twins	Full stack of KER1 metrics R1–R11	Project-based funding, strong scientific visibility	Strategic research field

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.

4.3.4 Identification of key players in the value chain

The following table summarises key archetypes and example organisations relevant for KER1. It is not exhaustive but illustrates the diversity of actors.

Table 6. Key players in the KER1 value chain

Role in value chain	Type of organisation	Example players	Relevance to KER1
End customers - pharma & biotech	Large and mid-size pharma with MASLD/liver or cardio-metabolic pipelines	Novo Nordisk, Eli Lilly, Gilead, BMS, Pfizer, Boehringer Ingelheim	Sponsors of MASLD/MASH and CLD trials, buyers of RUO quantitative biomarker services and multimodal analytics
End customers - CROs & imaging core labs	Global CROs and specialist imaging providers	ICON (including former Bioclinica), IQVIA, Labcorp, Median Technologies[23], IXICO	Intermediate buyers who integrate KER1 pipelines into their service offering for sponsors
End customers - clinical research consortia	EU-funded and national consortia, registries, hospital networks	EASL-affiliated liver cohorts, OHDSI-Europe, EHDEN nodes	Deploy KER1 in TREs for harmonised imaging/omics processing and VHT-supportive datasets
End customers - tertiary hepatology centres	University hospitals and specialised centres	Large referral centres in EU capitals, transplant centres	Long-term market for KER1-derived SaMD or advanced quantitative reporting tools
Upstream providers - imaging hardware	Scanner manufacturers (CT/MRI/US)	Siemens Healthineers, GE HealthCare, Philips, Canon Medical	Provide raw imaging data, own post-processing platforms, potential partners or competitors depending on integration strategy
Upstream providers - digital pathology	Slide scanners & pathology IT	Leica Biosystems, Philips Digital Pathology, Roche Ventana	Provide digital slides and LIS integration needed for R1/R8 pipelines
Upstream providers - omics	Mass spectrometry & NMR vendors, omics service labs	Bruker (NMR), Thermo Fisher, service providers	Underpin R10/R11 workflows, potential co-marketing partners
Horizontal platforms - multimodal AI/TRE	Cloud-based platforms integrating multimodal health data	SOPHiA GENETICS (DDM), DNAnexus, Flywheel	Potential B2B channels for embedding KER1 pipelines as analysis apps
Horizontal platforms - hospital IT	EHR, PACS, CDMS vendors	Epic, Oracle Cerner, Sectra, GE, Philips, Medidata	Gatekeepers for integration in routine care and trials, relevant for SaMD and CDMS integration

[23] https://mediantechnologies.com/clinical-trials

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.

Direct competitors - AI liver imaging & QIBs	Specialised liver imaging companies	Perspectum (LiverMultiScan), other QIB providers	Offer validated liver MRI biomarkers and services competing with some R2–R6 applications
Direct competitors - AI digital pathology for MASH	Digital pathology AI vendors	PathAI (AIM-MASH), HistoIndex, Aignostics	Provide AI-based histology scoring for MASH and fibrosis, overlapping with R1/R8 use cases
Indirect competitors - radiology AI suites	Vendor-specific AI post-processing	Siemens AI-Rad Companion, GE Edison AI, Philips IntelliSpace AI	Offer organ volumetry and other quantitative outputs, do not typically expose low-level features or open pipelines but may limit demand for separate tools in some settings
Substitute solutions - open source / in-house	Research frameworks and local scripts	3D Slicer, ITK-SNAP, MONAI, in-house pipelines	Allow hospitals and consortia to build their own pipelines, KER1 must compete on robustness, documentation and integration rather than on pure functionality
Regulators & policy	Regulatory and policy bodies	EMA, FDA, national HTAs, EDITH/VHT initiatives	Influence acceptance of model-based endpoints and AI tooling, KER1 must align with emerging guidance on QIBs, AI and VHTs[24]

This mapping highlights that KER1 competes most directly with **s**pecialised liver imaging and digital pathology AI vendors and, to a lesser extent, with general radiology AI suites in hospital settings. In trials and research, KER1's modular, multi-scale scope and openness offer a differentiating angle.

[24] www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/artificial-intelligence

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.

4.3.5 Relevant products and comparative analysis

A subset of products is particularly relevant when positioning KER1. The following table summarises representative examples and contrasts them with KER1's envisaged offering.

Table 7. Relevant products and comparative analysis with KER1

Product	Vendor	Main function	Target segment	Comparison with KER1
AIM-MASH / AIM-NASH	PathAI	AI-based tool for scoring MASH histology (CRN-based), qualified by EMA to assist a single central pathologist in phase 2/3 MASH trials and under FDA review.[25]	Pharma/CRO liver trials	Highly focused, regulated histology endpoint solution, optimised for enrolment and outcome assessment. KER1's pathology components (R1, R8, R9) are broader (multiple markers, microarchitecture) and more research-oriented at this stage, they support mechanistic modelling and multimodal signatures rather than a single regulatory endpoint.
LiverMulti-Scan	Perspectum	Multiparametric MRI solution providing non-invasive, standardised quantitative metrics (e.g. cT1, PDFF) for liver health, CE-marked and FDA-cleared, used in clinical practice and trials[26].	Clinical diagnostics, pharma trials	Strong focus on clinically validated MRI biomarkers and a streamlined clinical workflow. KER1's imaging stack (R2–R6) covers similar metrics but with more flexible, modular pipelines aimed at multicentre trials and modelling environments, clinical SaMD use will require additional development and validation.
AI-Rad Companion (liver and abdominal modules)	Siemens Healthineers	Suite of AI-powered radiology tools providing automated organ segmentation, volumetry and selected quantitative measures integrated into the vendor's imaging and PACS ecosystem.	Radiology departments	Tight integration with Siemens scanners and workflow, limited openness of feature-level outputs. KER1 does not compete at the level of radiologist UI, but can provide richer, model-ready features in multicentre and cross-vendor contexts, often sitting behind the scenes in TREs or platforms.

[25] https://www.ema.europa.eu/en/news/ema-qualifies-first-artificial-intelligence-tool-diagnose-inflammatory-liver-disease-mash-biopsy-samples

[26] https://www.perspectum.com/our-products/livermultiscan

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.

SOPHiA DDM™ for Multimodal[27]	SOPHiA GENETICS	Multimodal data platform combining genomic, imaging and clinical data, with ML analytics and growing support for multimodal signatures and patient stratification.	Pharma and academic precision medicine programmes	Acts as a platform for multimodal analytics, not specialised in liver or in multiscale mechanistic modelling. KER1 can complement such platforms by providing liver-specific, model-aligned pipelines integrated as Docker "apps" or services.
Clinical trial imaging services	Median Technologies, IXICO[28], others	Central imaging services (acquisition protocol design, image management, endpoint analysis, including some AI/radiomics)	Pharma/CRO trials	These providers often build bespoke pipelines per study, some are diversifying into AI. KER1 can be licensed or white-labelled as a standardised pipeline suite embedding domain adaptation and harmonisation, or it could compete as a specialised liver QIB provider in focused niches.
Open-source frameworks (3D Slicer, MONAI)	Community projects	Generic imaging analysis and deep learning frameworks enabling segmentation and radiomics via reusable components	Hospitals, research labs	These are powerful but require local expertise for configuration and QA. KER1 differentiates by offering curated, validated, liver-specific, end-to-end pipelines with explicit model-alignment and documentation rather than generic building blocks.

Overall, KER1's competitive angle is not to outperform every specialised product in its own narrow niche, but to provide an integrated, liver-centric, multimodal processing layer that:

- spans imaging, digital pathology and omics,
- is explicitly designed to feed multiscale mechanistic and AI models,
- incorporates cross-site harmonisation and domain adaptation,
- can be embedded in multiple environments (CROs, consortia, platforms).

This system-level positioning reduces head-to-head competition with single-function SaMDs while still allowing specific components (e.g. fat fraction pipelines, radiomics) to be commercialised in more conventional ways when appropriate.

[27] https://www.sophiagenetics.com/sophia-ddm-for-multimodal

[28] https://ixico.com/services/neuroimaging-biomarkers

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.

4.3.6 Market dynamics and value network

The adoption of KER1-type solutions is governed by several key market dynamics, barriers, and enablers.

Procurement Processes: This dynamic strongly favours the pharma/CRO market. Hospital procurement is notoriously slow, complex, IT-department-led, and focused on capital budgets and workflow integration[29]. In contrast, Pharma R&D procurement is value-based and can be much faster for solutions that demonstrably accelerate R&D, often funded from operational budgets[30].

Data Governance & Interoperability: This is the central challenge and opportunity. Data is siloed, and multicentre sharing is a legal and technical minefield due to GDPR and local site rules[31]. This dynamic is a *barrier* to naive data-pooling but a massive *driver* for KER1's federated-ready, containerised design, which can "bring the compute to the data." The rise of TREs[32] and federated networks (EHDEN[33]) is the dominant architectural trend.

Validation & Trust: A major barrier to AI adoption in healthcare is a lack of trust and a fear of "black-box" systems[34]. Clinicians and researchers demand transparency and extensive real-world validation of clinical endpoint achievement[35]. KER1's origins in an open, academic research project (ARTEMIS) and its adherence to public standards (e.g., IBSI) are key *enablers* to build this trust.

Competition & Vendor Lock-in: Incumbent imaging vendors (like Siemens) leverage their market power to create closed ecosystems, locking customers into their PACS and AI platforms. The market is also consolidating. KER1's *open, modular, and vendor-agnostic* nature is a key differentiator for customers (AMCs, platforms, pharma) who want to avoid this lock-in.

Table 8. Barriers and enablers for KER1 adoption

Factor	Barrier or Enabler?	Impact on KER1 Adoption	Possible Mitigation or Leverage Strategy
European Health Data Space (EHDS)	**Enabler (Massive)**	Creates the legal and technical framework for secondary use of health data[36], generating massive, EU-wide demand for	**Leverage:** Position KER1 as the *de facto* "processing engine" for EHDS data access bodies[37] to make raw,

[29] Barriers to AI in Radiology: Why Clinical Adoption Is Still Struggling? https://www.deepc.ai/blog/barriers-to-ai-in-radiology-why-clinical-adoption-is-still-struggling

[30] Clinical Research SaaS Technology: 10 Powerful Ways for Faster Trials - Lifebit, https://lifebit.ai/blog/clinical-research-saas-technology

[31] Challenges related to data protection in clinical research before and during the COVID-19 pandemic: An exploratory study - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC9589288/

[32] Trusted Research Environments (TRE) Platform | DNAnexus®, https://www.dnanexus.com/trusted-research-environment-platform

[33] Enabling Large-Scale Federated Data Analysis in Europe through the OMOP Common - OHDSI, https://www.ohdsi.org/wp-content/uploads/2018/10/Global-Panel-All-Slides.pdf

[34] Barriers to AI in Radiology: Why Clinical Adoption Is Still Struggling? https://www.deepc.ai/blog/barriers-to-ai-in-radiology-why-clinical-adoption-is-still-struggling

[35] Medicine, healthcare and the AI act: gaps, challenges and future implications https://pmc.ncbi.nlm.nih.gov/articles/PMC12282355/

[36] European Health Data Space (EHDS) https://www.european-health-data-space.com/

[37] Reuse of health data - Public Health - European Commission, https://health.ec.europa.eu/ehealth-digital-health-and-care/reuse-health-data_en

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.

		harmonization/quantification tools.	heterogeneous data research-ready[38]
VHT & EU Strategic Initiatives	**Enabler**	Aligns KER1 with high-level EU strategic goals (VHT, Digital Europe)[39], providing access to non-dilutive funding, validation partners, and a "blue ocean" market.	**Leverage:** Actively partner with EDITH-CSA and VHT projects to become the reference implementation for the VHT data-input layer.
Data Governance & GDPR	**Barrier & Enabler**	*Barrier:* Complexity of multicenter data sharing and divergent national rules. *Enabler:* Drives demand for robust, traceable, federated pipelines (which KER1 is).	**Mitigate/Leverage:** Deploy KER1 within partner TREs or on-premise, leveraging its containerised design to "bring the compute to the data."
Lack of Clinician / User Trust	**Barrier**	Widespread skepticism of AI/quantitative tools without extensive real-world validation. "Black-box" fear.	**Mitigate:** 1) Pursue an "open validation" strategy: publish results, validate against standards (IBSI), and partner with research consortia to build a body of evidence. 2) Apply usability engineering (EN 62366) to consider end-user requirements and validation to enhance trustworthiness.
Integration with PACS/EHR	**Barrier**	Deep workflow integration is the n°1 hurdle for clinical AI adoption. Requires partnerships with incumbent vendors who have closed systems.	**Mitigate: Bypass this barrier initially.** Focus on the *offline* research/trial markets (Pharma, Consortia) first, which do not require real-time PACS integration.
Competition (Incumbents & Platforms)	**Barrier**	Incumbent (Siemens) and platform (SOPHIA) competitors offer end-to-end solutions. The market is consolidating.	**Mitigate: Differentiate.** KER1 is not another "diagnostic AI" or "data host." It is a *modular, open, VHT-ready "data refinery"* that can *partner* with platforms (Flywheel)
MASLD Market Need (QIBs)	**Enabler (Major)**	Urgent, high-value commercial need for non-invasive biomarkers in pharma trials, which are booming.[6]	**Leverage:** Position KER1's MASLD-specific pipelines (R1, R4, R8, R10, R11) as the base product to target the Pharma/CRO market.

[38] The European Health Data Space (EHDS): The Promise of Secondary Use of Data for Healthcare Innovation - MedTech Europe, https://www.medtecheurope.org/medtech-views/policy-views/the-european-health-data-space-ehds-the-promise-of-secondary-use-of-data-for-healthcare-innovation/

[39] Virtual Human Twins initiative for health and care https://digital-strategy.ec.europa.eu/en/policies/virtual-human-twins

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.

5 KER2 – Multi-modal clustering of patients and pattern identification

5.1 Functional analysis

The development of *"Multi-modal clustering of patients and pattern identification"* pipelines, which corresponds to KER 2 within ARTEMIs, is motivated by critical overarching scientific, clinical, and market needs in translational research and personalized medicine. These pipelines address the complexity inherent in high-volume, heterogeneous patient data by enabling structured data mining for discovery and enhanced prediction.

The main applications and scientific needs motivating this development include:

5.1.1 Enabling Patient Stratification and Personalized Management

The primary market and clinical need is to move beyond "average patient" treatment strategies by identifying specific patient subgroups, or phenotypes, within a large cohort. Clustering methods allow big data specialists to employ AI-based approaches to identify patient subgroups sharing similar disease states or quantitative characteristics, such as specific blood measurements, patient specific characteristics (e.g., age, weight), and omics profiles (proteomic and metabolomic). This capability is foundational for refining "generic Virtual Twins" into "patient-specific Virtual Twins" by tailoring sub-models to specific subpopulations, thereby enhancing the precision of therapeutic strategies.

5.1.2 Generating Mechanistic Hypotheses and New Knowledge

Clustering and pattern identification are essential scientific tools for transforming large population data into actionable knowledge. The analysis of identified patterns can generate new mechanistic hypotheses for clinical experts to assess, which are then formalized and explored using mechanistic models (WP5 and WP6). This process systematically identifies relationships and mechanisms (T4.3) that contribute to enhanced knowledge of complex disease progression mechanisms, such as fibrosis or portal hypertension.

5.1.3 Bridging Multi-Scale and Multi-Domain Data Gaps

Modern clinical research involves multimodal data (clinical records, medical imaging, histology, and omics profiles). Clustering pipelines provide the framework to cope with the "heterogeneity of the federated cohorts" and holistically integrate these disparate data features. A key application is establishing systematic relationships between non-invasive markers (blood biomarkers, medical imaging) and invasive, histological findings, thereby inferring information that could potentially avoid biopsies.

5.1.4 Informing and Calibrating Computational Models

The patterns and clusters identified directly serve as critical input for downstream computational models. This data is used as "patient group specific features" to enhance the precision of Ordinary Differential Equation (ODE) models and deep learning models. Furthermore, pattern identification contributes to the continuous refinement and validation of Machine Learning models, ensuring their trustworthiness and robustness by identifying and transparently disclosing uncertainties and possible biases in prediction scenarios.

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.

5.1.5 Supporting Drug and Intervention Evaluation

In a broader application, clustering assists in tackling complex clinical questions related to treatment efficacy, such as evaluating the impact of specific drugs (like SGLT2 inhibitors or statins) or interventions (like bariatric surgery or TIPS placement) on disease progression (e.g., liver fibrosis and cardiovascular comorbidities). These data-driven analyses inform modelers about drug-specific effects and relevant correlations.

5.2 Technology Analysis

5.2.1 Limitations of the current state of the art

Over the last few years, multimodal machine learning and patient stratification have advanced substantially. Large MASLD cohorts now support robust clustering of patients with different long-term risks, as shown by recent data-driven analyses combining clinical variables with histology and genetics across ABOS, UK Biobank and other cohorts[40]. More broadly, reviews of multimodal machine learning in precision health and chronic liver disease highlight a fast-growing ecosystem of methods and data resources[41]. Commercial platforms such as Tempus Lens[42] in oncology and multimodal AI solutions from Owkin and Siemens Healthineers' AI-Rad Companion also demonstrate that integrating heterogeneous clinical, imaging and molecular data is technically feasible at scale in selected domains.

However, when viewed from the perspective of ARTEMIS and KER2, several critical gaps remain that limit the systematic use of multimodal clustering and pattern identification as inputs to virtual-twin-type models in chronic liver disease and MASLD.

5.2.1.1 Data infrastructure and cohort scope

Recent MASLD clustering work relies on large, but still highly curated, research cohorts where inclusion criteria, follow-up schedules and data modalities are carefully controlled[43]. In contrast, many mechanistic and predictive models in hepatology and cardiometabolic disease are still developed on relatively small, monocentric or trial-derived datasets with limited ethnic diversity and incomplete capture of comorbidities. Reviews of AI in MASLD and alcohol-related liver disease emphasise that most published models use thousands rather than hundreds of thousands of patients and often lack external validation or representation of real-world clinical heterogeneity[44]. Even when large observational resources such as UK Biobank or EHR-based cohorts are used, the subsets with deep liver phenotyping (imaging biomarkers, histology, omics) remain comparatively small and are not readily reusable as harmonised "plug-and-play" resources for clustering across centres[45].

As a result, multi-modal clustering is typically implemented within individual research groups for specific studies, rather than as generalisable pipelines that can be deployed across federated, heterogeneous clinical cohorts of MASLD patients.

[40] https://www.nature.com/articles/s41591-024-03283-1
[41] https://www.nature.com/articles/s41746-022-00712-8
[42] https://www.tempus.com/?srsltid=AfmBOoox1QnxSAn1niwzTzbJfCwsJkuRb7N1Tjx7I88ojZbSVBXoHf4
[43] https://www.nature.com/articles/s41591-024-03283-1
[44] https://www.oaepublish.com/articles/mtod.2024.84
[45] https://www.ejrai.com/article/S3050-5771%2825%2900014-3/fulltext

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.

5.2.1.2 Integration and harmonisation of heterogeneous modalities

State-of-the-art multimodal models in precision health commonly fuse two modalities (e.g. imaging plus EHR, or imaging plus genomics), and even in ambitious applications the integration is often limited to a subset of the available data types[46]. In MASLD and chronic liver disease, AI-guided integration of omics and clinical data is recognized as promising, but reviews consistently report major obstacles: non-standardised pre-processing, batch effects across centres, missing-data patterns that are modality-specific, and a lack of reproducible, open pipelines for end-to-end data fusion[47].

Moreover, many published works treat imaging, histology, omics and routine clinical data in separate analytic workflows, with manual, ad hoc steps to combine outputs. This separation complicates the construction of robust, quantitative feature spaces that can be used consistently for unsupervised clustering across federated cohorts, and it hinders the definition of standardised "multi-modal fingerprints" that could be shared with mechanistic model developers.

5.2.1.3 Immature descriptive clusters

Recent MASLD cluster analyses demonstrate that purely data-driven phenotypes can stratify risks of incident chronic liver disease, cardiovascular events and diabetes across large populations[48]. Nonetheless, most of these approaches remain fundamentally descriptive: clusters are characterised by risk profiles, but are only loosely connected to underlying mechanistic hypotheses or to the structure of virtual-twin models. Reviews of multi-omics integration in chronic liver disease similarly highlight that, while molecular signatures can distinguish disease stages or progression trajectories, the mapping from high-dimensional patterns to biologically interpretable pathways and model parameters is still immature[49].

This limits the possibility of using clusters as "model-ready abstractions", for example, to define subgroup-specific parameter ensembles, to identify non-invasive surrogates for histological features, or to support systematic generation and testing of mechanistic hypotheses within digital-twin frameworks. Scoping reviews of human digital twins in healthcare further show that only a small fraction of existing models meet stringent criteria for truly patient-specific, dynamically updated twins with quantified uncertainty, and that liver disease remains under-represented compared with cardiovascular and oncology applications[50].

5.2.1.4 Lack of federated and privacy-preserving multimodal clustering

Federated learning has matured rapidly in medical image analysis and EHR-based prediction, but the vast majority of deployed or prototyped systems focus on supervised tasks such as classification, segmentation or outcome prediction[51]. Surveys of federated learning on multimodal data underline additional challenges (statistical heterogeneity between sites, unbalanced modality availability, communication constraints and the difficulty of aligning feature spaces across institutions) and note that unsupervised clustering methods tailored to such

[46] https://www.nature.com/articles/s41746-022-00712-8

[47] https://pmc.ncbi.nlm.nih.gov/articles/PMC11874365

[48] https://www.nature.com/articles/s41591-024-03283-1

[49] https://www.researchgate.net/publication/362501132_Multi-Omics_and_Artificial_Intelligence-Guided_Data_Integration_in_Chronic_Liver_Disease_Prospects_and_Challenges_for_Precision_Medicine

[50] https://www.nature.com/articles/s41746-025-01910-w

[51] https://www.researchgate.net/publication/394560001_Federated_Learning_for_Medical_Image_Analysis_Privacy-Preserving_Paradigms_and_Clinical_Challenges

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.

settings are still rare and largely confined to method papers or small-scale demonstrations[52].

In parallel, conceptual work on multimodal diagnostics highlights federated learning as a future direction rather than an established solution, stressing that current deployments remain limited, centrally hosted, and focused on relatively homogeneous data sources[53]. For MASLD and chronic liver disease specifically, there are no widely adopted, off-the-shelf federated clustering frameworks that can operate on longitudinal, multi-modal data (clinical, imaging, histology, omics) distributed across multiple hospitals while preserving privacy and regulatory compliance.

5.2.1.5 Tooling, interpretability and clinical adoption

A growing number of commercial platforms now offer multimodal data integration and patient stratification services, particularly in oncology. Tempus, for example, provides multimodal real-world data and analytics via its Lens platform, combining structured clinical data, genomics and imaging to support cohort building and trial optimisation[54]. Owkin develops multimodal AI models and federated learning infrastructures for drug development and clinical trials, mainly in oncology and pathology[55]. Large imaging vendors such as Siemens Healthineers deploy multi-modality imaging decision support (AI-Rad Companion) and pathway-oriented platforms that integrate imaging, selected clinical data and workflow information[56].

Despite this progress, several limitations persist from the viewpoint of KER2. First, most commercial solutions are closed ecosystems oriented to specific diseases (mostly cancer) and do not expose generalisable, transparent clustering pipelines that could be repurposed for MASLD or tightly coupled to mechanistic virtual-twin models. Second, interpretability and uncertainty quantification for multimodal patient stratification are still underdeveloped, with many models functioning as "black boxes", systematic frameworks for explaining clusters, assessing stability and communicating uncertainty to clinicians remain the exception rather than the rule[57]. Third, reviews of multimodal AI in medicine and of human digital twins underline that very few multimodal tools have reached regulatory approval or routine clinical use, and that gaps in interoperability, reporting standards and validation continue to slow adoption[58].

Taken together, these limitations indicate that, despite rapid advances and a rich ecosystem of methods and products, there is still no mature, open and federated framework that (i) robustly harmonises heterogeneous liver-related data across centres, (ii) performs clinically meaningful multimodal clustering and pattern identification, and (iii) produces interpretable, model-ready outputs that can drive and calibrate virtual-twin-type models in MASLD and chronic liver disease. Addressing this gap is the core motivation for the development of KER2 within ARTEMIS.

5.2.2 ARTEMIS results and technical features beyond the state of the art

[52] https://link.springer.com/article/10.1007/s11633-022-1398-0

[53] https://www.researchgate.net/publication/392708497_Artificial_Intelligence_in_Multimodal_Diagnostics_Integrating_Imaging_Genomics_and_EHRs_for_Precision_Medicine

[54] https://www.tempus.com/?srsltid=AfmBOoox1QnxSAn1niwzTzbJfCwsJkuRb7N1Tjx7I88ojZbSVBXoHf4

[55] https://www.owkin.com/newsfeed/ai-and-multimodal-patient-data-for-better-faster-clinical-trials

[56] https://www.siemens-healthineers.com/en-th/digital-health-solutions/digital-solutions-overview/clinical-decision-support/ai-rad-companion

[57] https://www.mdpi.com/2078-2489/16/11/971

[58] https://www.nature.com/articles/s41746-022-00712-8

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 9..Main results clustered under KER2

ID	WP Task(s)	Main partners contributing	Result / component	Description and functionality	Technical features beyond the state of the art	Exploitation potential and link to KER2
R1	T4.1	BU (lead), ICAN, HULAFE, MUV, ULEI, INRIA, MEDEX, others	**Federated clinical clustering pipeline for fibrosis risk stratification**	Unsupervised clustering pipeline that uncovers hidden subgroups among chronic liver disease patients based on, routine clinical variables, fibrosis risk scores (FIB-4, NFS, APRI, AST/ALT), demographics and comorbidities. Initially developed on public biopsy-proven NASH data (~215 patients) and being adapted to the federated ARTEMIS cohort. It provides risk-oriented stratification (e.g. advanced fibrosis vs non-advanced) and cluster-specific profiles to support mechanistic and AI model parameterisation.	Development and early validation of scalable, privacy-guaranteed federated clustering tailored to longitudinal, non-IID clinical data using the Flower framework[59]. Goes beyond single-centre, centrally pooled clustering by enabling horizontal federated experiments, and targets non-parametric clustering in future iterations to capture natural groupings in CLD cohorts.	Core federated patient-stratification module in KER2. The resulting clusters act as "patient-group specific features" that refine generic Virtual Twins into patient-specific Virtual Twins and support risk-stratified evaluation of drugs and interventions (e.g. SGLT2i, statins, bariatric surgery, TIPS). In the medium term, it can be packaged as a federated analytics component for hospital platforms or clinical decision-support systems.
R2	T4.1	HULAFE (lead), input from BU / WP3 imaging pipelines	**Uncertainty-aware voxel-wise imaging clustering for liver MRI**	Clustering method applied directly to dynamic multi-echo liver MRI sequences and other selected MRI series from WP3. Repeated sampling with added noise is used to obtain voxel-	Voxel-wise clustering with explicit uncertainty estimation, moving beyond traditional radiomics extraction. Each voxel classification is accompanied	Provides structured quantitative imaging information derived from T3.2/T3.3 pipelines (fat fraction, stiffness, radiomics features) for integration in

[59] https://flower.ai

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.

				wise cluster assignments, generating tissue-type maps and associated confidence measures. The method extracts spatially resolved imaging biomarkers (and their distributions) and enables correlation with fibrosis stage and related comorbidities.	by a confidence metric, allowing robust tracking of regional changes over time and more transparent imaging-based pattern identification. This directly addresses the SOTA gap in systematic, quantitative links between non-invasive imaging markers and histological or clinical endpoints.	the global multimodal clustering framework in T4.3. Within KER2, it supplies a non-invasive imaging layer that can be combined with clinical and omics data to derive patterns potentially allowing biopsy-sparing surrogates and imaging-driven Virtual Twin calibration.
R3	T4.2	BU, ICAN, HULAFE, others	**Omics-based phenotype profiling and biomarker discovery pipeline**	Omics analysis pipeline for NASH/CLD phenotypes combining differential gene expression (DEGs), protein–protein interaction (PPI) networks and KEGG pathway enrichment. Public datasets from GEO were used to compare healthy vs NASH samples and to generate PPI networks and pathway profiles, the same workflow is prepared to ingest ARTEMIS proteomic and metabolomic data from WP3. It outputs phenotype profiles, hub-gene panels and pathway signatures that	Uses advanced graph-theory metrics (betweenness and degree centrality, k-degenerate/DMNC) and other ML methods to identify key genes at the interface of NASH progression, including metabolic risk factors. This goes beyond simple DEG lists by extracting multi-disease "common DEGs" and pathway-level mechanisms that can be used as mechanistically interpretable features for clustering and modelling.	Supplies a reusable bioinformatics workflow that feeds omics-derived features into KER2 clustering and into mechanistic hypothesis generation in WP5 (e.g. intracellular metabolic models, signalling pathways). It has clear potential as a service or software module for biomarker discovery, patient stratification in trials, and drug repurposing, all anchored in the multi-modal KER2 framework.

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.

				characterise distinct molecular subgroups.		
R4	T4.3	BU, HULAFE, others	**Deep federated multi-modal clustering pipeline for prognostic subgroups**	Multi-modal data analysis pipeline, initially validated on the WAW-TACE public dataset of HCC patients, to identify prognostic subgroups undergoing trans-arterial chemoembolization (TACE). Combines clinical trajectories, treatment/outcome data and radiomics features from CT/MRI. Uses deep autoencoder-based representation learning within a horizontal federated clustering architecture, followed by evaluation of cluster robustness and prognostic value (t-SNE/UMAP visualisation, Kaplan–Meier curves, Cox models).	Implements a deep federated clustering framework where clients communicate autoencoder weights and local centroids to a central server that returns global cluster assignments (pseudo-labels) under privacy constraints. Focuses on model-based probabilistic approaches (e.g. Dirichlet process / Gaussian mixture) and explores two-mode clustering (groups of patients and groups of variables), which is rarely addressed in current clinical multi-modal AI pipelines.	Constitutes a central software asset of KER2, demonstrating how multi-modal clinical + imaging + omics features can be fused in a federated, privacy-preserving manner to uncover prognostic subgroups. It is directly reusable as a generic multi-modal analytics engine for other indications and cohorts and can be further developed into a clinical decision-support or trial-enrichment module built on top of KER2.
R5	Cross-cutting WP4 outcome	BU, HULAFE, ICAN, ULEI, INRIA, MEDEX, others (WP3/WP5 interfaces)	**Integrated KER2 pattern repository and interface to mechanistic models**	Consolidated repository that organises patient clusters, prognostic subgroups, imaging-derived patterns and omics-based phenotypes into structured “pattern objects” and patient-group profiles. It aggregates outputs from	Provides an explicit pattern layer tightly coupled to mechanistic models, addressing a key SOTA gap: tools that translate data-driven clusters into formal mechanistic hypotheses and calibration scenarios. By	This integrated repository is the operational embodiment of KER2. It underpins the refinement of Virtual Twins, supports systematic exploration of non-invasive vs invasive relationships (e.g. imaging vs histology), and

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.

		clinical/federated clustering (KER2.1), imaging clustering (KER2.2), omics profiling (KER2.3) and deep multi-modal clustering (KER2.4), and exposes them as model-ready inputs (cluster labels, centroids, feature importances, uncertainty metrics, survival-linked signatures) for WP5–WP6 Virtual Twins.	integrating federated learning, uncertainty-aware imaging clustering and mechanistically interpretable omics signatures, it creates a unique bridge between AI-driven discovery and explainable digital twins in MASLD/CLD.	enables Virtual-Twin-based services such as treatment scenario simulation, drug-effect analysis and trial stratification per cluster. In future exploitation, it can serve as a premium analytics and "pattern API" layer on top of the ARTEMIS platform for hospitals, CROs and industry partners.

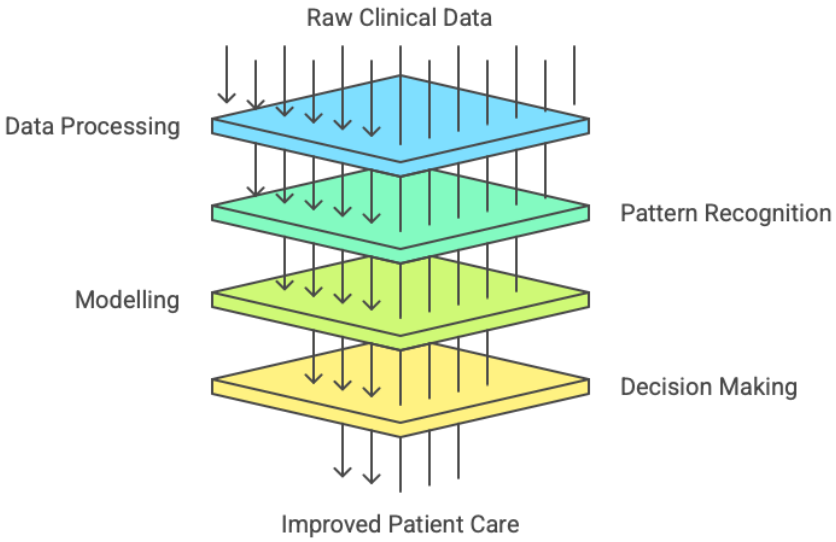
This document is a deliverable of the ARTEMIs project. The ARTEMIs Project is funded by the European Union within the Horizon Europe program under grant agreement No 101136299.

5.3 Market analysis

5.3.1 Value chain overview

The value chain surrounding KER2's multi-modal analytics pipeline begins with **Data Processing and Quantification**. The primary actors at this stage are imaging core labs, hospital IT departments, and specialized vendors like Siemens Healthineers or Perspectum. Their core activity involves transforming raw, heterogeneous clinical data, spanning imaging, digital pathology, and omics, into standardized quantitative metrics. This process results in an output of structured, model-ready features. However, this output is often characterized by high dimensionality and a lack of integrated biological interpretation, requiring advanced analytics to bridge the gap between isolated metrics and patient-level understanding.

The second link in the value chain is **Pattern Recognition and Patient Stratification**, which is the layer where KER2 creates most of its value. Its core activity involves ingesting the processed multi-modal metrics and applying federated clustering algorithms to identify latent patient subgroups and complex risk profiles. This process results in an output of robust, privacy-preserved patient phenotypes and predictive biomarker patterns. This output is characterized by its ability to resolve heterogeneity that single modalities cannot reveal, differentiating it from single-point solutions like PathAI by offering a holistic view of disease progression.



The third link in the value chain is **Modelling and Insight Generation**. The primary actors at this stage are computational biologists, Virtual Human Twin developers, and regulatory science teams involved in initiatives like EDITH. Their core activity involves consuming the stratified clusters from KER2 to calibrate, train, and refine mechanistic models and in-silico simulations. This process results in an output of "digital twins" of patient subpopulations, characterized by enhanced biological realism and predictive power for specific disease trajectories.

The final link in the value chain is **Decision and Use**. The primary actors at this stage are the end-users, including pharmaceutical R&D directors at companies like Boehringer Ingelheim, CROs, and specialized hepatologists. Their core activity involves applying the stratified insights and simulation results to critical challenges, such as enriching clinical trials with high-risk patients or tailoring personalized treatment plans. This process results in an output of data-driven decisions that de-risk drug development and improve patient care, providing a critical feedback loop to refine the upstream KER2 clustering parameters.

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.

5.3.2 Potential Application Fields for KER2

Given ARTEMIS's focus on liver and cardio-metabolic diseases, the most relevant application fields for KER2 are those where multi-modal patient stratification can resolve heterogeneity, reduce trial failure rates, or enable precision medicine. Below we summarize these fields and narrow down to high-potential segments for Europe.

5.3.2.1 Pharma / CRO Clinical Trials in MASLD/MASH and Cardio-Metabolic Disease

The most attractive application field for KER2 is optimizing patient selection and endpoint analysis for pharma and CRO-run clinical trials. In MASLD/MASH trials, high patient heterogeneity often impedes efficacy signals, and invasive biopsies remain a bottleneck. Regulators and sponsors are actively seeking AI tools to reduce variability and identify likely responders.

KER2's clustering capabilities map directly to high-value needs in this segment: 1) stratifying patients by fibrosis risk or co-morbid profiles to tailor inclusion criteria, 2) discovering novel composite biomarkers that combine imaging and omics, and 3) enriching trials with high-risk phenotypes to improve statistical power.

Demand is strong, with over 1,200 interventional MASLD/MASH trials launched globally by 2024. The AI-enabled clinical trials market is valued at approximately $2.6 billion in 2025 and is projected to grow at roughly 27% annually through 2034, according to Precedence Research[60]. AI solutions in this space have reportedly improved enrolment rates by 11-20%, as highlighted by Toronto-based startup Altis Labs[61]. For KER2, the immediate opportunity is offering "research-as-a-service" analytics to sponsors like Boehringer Ingelheim or Novo Nordisk, similar to how Tempus leverages its multimodal data in strategic collaborations for cancer pipelines[62].

5.3.2.2 Multi-Centre Research Consortia and Registries

A second key field is large-scale academic and clinical research networks (e.g., EASL networks, OHDSI-Europe, EHDEN) amassing real-world data on liver and metabolic conditions. Their bottleneck is not data volume, but the lack of robust pipelines to harmonize heterogeneous data and extract meaningful patient phenotypes across sites.

KER2 directly addresses this gap by deploying as a containerized analysis layer within *Trusted Research Environments (TREs)*. It performs federated clustering on multi-site data without pooling sensitive records, yielding reproducible phenotypes critical for longitudinal studies.

While individual budgets are modest compared to pharma, this segment represents a sustained opportunity estimated at ~€100 million/year in the EU for imaging/omics-intensive data processing. Market drivers such as the demand for efficient data management in trials continue to grow, with the global AI in clinical trials market expanding rapidly[63]. Moreover, the *European Health Data Space (EHDS)*[64] will drive massive demand for such harmonization tools to enable secondary use of health data.

---

[60] https://www.precedenceresearch.com/ai-in-clinical-trials-market

[61] LinkedIn Pulse analysis of AI clinical research companies https://www.linkedin.com/pulse/ten-ai-companies-clinical-research-watch-2023-andrii-buvailo/

[62] https://www.tempus.com/news/tempus-enters-multi-year-strategic-collaboration-with-boehringer-ingelheim-to-advance-its-cancer-pipeline/

[63] https://www.marketsandmarkets.com/Market-Reports/ai-in-clinical-trials-market-42687548.html

[64] health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en

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.

5.3.2.3 Commercial Multimodal Data & AI Platforms (B2B Channel)

A third field is a channel strategy: integrating KER2 into existing cloud-based platforms that serve biomedical researchers, such as DNAnexus, Flywheel, or SOPHIA GENETICS. These platforms cater to pharma and academia but often lack specialized liver/metabolic analytics.

KER2 could be embedded as a premium "pattern discovery" module (via Docker or API) within these platforms. This B2B approach leverages the partners' existing salesforce and installed base. For instance, in 2021, GE HealthCare partnered with SOPHIA GENETICS to match treatments to multimodal data[65]. By positioning KER2 as a compatible plugin for platforms like SOPHIA DDM, the project can tap into a market segment where multimodal analysis is becoming standard.

5.3.2.4 Advanced Clinical Diagnostics in Tertiary Centres

In the longer term, KER2 can be deployed as a clinical decision support tool in specialized settings, such as tertiary hepatology or cardio-metabolic clinics. Here, KER2 would ingest patient data to assign "risk clusters" (e.g., high-risk fibro-inflammatory phenotype) to guide personalized therapy.

This segment offers high value but faces significant barriers, including the need for regulatory approval (MDR). However, the market is being primed: the AI in diagnostics market is expected to grow from $1.9 billion in 2025 to over $10 billion by 2034, as reported by Precedence Research[66]. Adoption in Europe is projected to follow a ~25% CAGR, according to Medical Economics[67], driven by the increasing need for precision tools in complex chronic diseases.

5.3.2.5 Virtual Human Twin (VHT) and In-Silico Trials

Aligned with ARTEMIS's vision, KER2 serves as a foundational technology for the emerging field of Virtual Human Twins. KER2's clusters provide the population parameters needed to create "digital twins" of patient subpopulations for in-silico experimentation.

Europe is heavily investing in this space through initiatives like the European Commission's VHT initiative[68]. While the commercial market for in-silico clinical trials is currently niche, it is growing as regulators begin developing frameworks for simulated evidence, with Grand View Research projecting steady expansion through 2030[69].

5.3.3 Prioritization and Narrowing of Target Markets

Considering the unmet needs, funding availability, and regulatory landscape, the segments with the highest potential for commercialization of KER2 by 2027 are:

1. **Pharma / CRO Clinical Trials:** Highest near-term revenue potential due to urgent need for patient stratification in failing NASH trials and available R&D budgets.
2. **Research Consortia & Registries:** Strong technical alignment with KER2's federated capabilities, creates necessary validation data and scientific visibility.

[65] https://www.gehealthcare.com/about/newsroom/press-releases/ge-healthcare-and-sophia-genetics-to-collaborate-to-match-treatments-to-multimodal

[66] https://www.precedenceresearch.com/artificial-intelligence-diagnostics-market

[67] https://www.medicaleconomics.com/view/the-state-of-ai-diagnostics-in-health-care-projected-24-6-cagr-through-2030

[68] https://digital-strategy.ec.europa.eu/en/policies/virtual-human-twins

[69] https://www.grandviewresearch.com/industry-analysis/insilico-clinical-trials-market-report

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. **Multimodal Platforms (B2B):** A strategic channel to scale adoption efficiently by piggybacking on existing biomedical data ecosystems.

Table 10. Analysis of market segments for KER2

Segment	Main unmet need	KER2 offer	Market characteristics	Priority (by 2027)
1. Pharma / CRO liver trials	High heterogeneity impedes efficacy signals, Need for non-invasive, data-driven endpoints to de-risk R&D.	Multi-modal patient stratification, Responder identification, Federated analysis for biomarker discovery.	Global AI trials market ~$2.6B (2025) [Precedence Research], High growth (~27%).	**High - Primary**
2. Multicentre consortia & registries	Lack of robust pipelines to harmonise heterogeneous multi-site data, GDPR constraints on data pooling.	Federated clustering, Privacy-preserving pattern mining, Harmonised phenotyping across sites.	EU niche opportunity ~€100M/year, Driven by EHDS and data sharing initiatives.	**High - Strategic & Validation**
3. Multimodal AI platforms (B2B)	Generic platforms lack specialized liver/metabolic analytics, Need for "apps" to enhance platform value.	Containerised "Pattern Discovery" modules (Docker/API) embedded in PaaS environments.	High growth (~35% CAGR) in AI analytics, Scalable via partnerships (e.g. [SOPHIA GENETICS]).	**Medium - Leverage Channel**
4. Advanced clinical diagnostics	Guidance for personalized therapy in complex cases (e.g. advanced MASH).	Clinical Decision Support (risk scoring, complex phenotyping) for tertiary care.	High potential (~$1.9B market) but high regulatory (MDR) barriers [Medical Economics].	**Low Short-term / Second Wave**
5. VHT / In-silico trials	Validated virtual cohorts for in-silico control arms, Calibration of mechanistic models.	Cluster-based population parameters, Digital twin cohort generation.	Emerging field, Strategic alignment with EC initiatives, Future regulatory acceptance [Grand View Research].	**Strategic Research Field**

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.

5.3.4 Identification of Key Players in the Value Chain

Bringing KER2 to market involves a complex network of stakeholders. The table below identifies key players, from end-users to competitors.

Table 11. Key players in the KER2 value chain

Role in Value Chain	Type of Organization	Example Players	Relevance to KER2
End Customers - Pharma & Biotech	Drug developers with pipelines in NASH, diabetes, cardio-metabolic.	Novo Nordisk, Eli Lilly, Gilead, Boehringer Ingelheim, Pfizer.	Sponsors of trials, buyers of stratification services to accelerate R&D and reduce failure rates [Tempus News].
End Customers - CROs & Core Labs	Clinical trial operations and specialized imaging/data services.	IQVIA, ICON, Labcorp, Median Technologies.	Intermediaries that can integrate KER2 into their service offerings to sponsors (e.g., AI-driven recruitment).
End Customers - Research Consortia	EU-funded networks, biobanks, and registries.	EASL networks, OHDSI-Europe, EHDEN, LITMUS.	Collective customers for federated clustering, key for validation and generating reproducible phenotypes.
Upstream Providers - Data Sources	Imaging vendors, Digital Pathology, Omics labs.	Siemens Healthineers, GE HealthCare, Philips, Roche Ventana.	Supply the raw multi-modal data KER2 inputs. Potential partnership opportunities for data standardization.
Horizontal Platforms (B2B)	Cloud-based health data & AI analytics platforms.	SOPHIA GENETICS, DNAnexus, Flywheel.	Channel partners to host KER2 as an app/module [SOPHIA GENETICS], reaching a broader user base.
Direct Competitors - Multimodal AI	AI-driven precision medicine companies.	Owkin (Oncology / Federated), Tempus (Precision Med), Unlearn.AI.	Compete on multi-modal insight generation. KER2 differentiates via liver focus and mechanistic model integration [LinkedIn Pulse].
Direct Competitors - Niche AI Tools	Domain-specific AI for liver imaging/pathology.	Perspectum (LiverMultiScan), PathAI (AIM-NASH).	Compete on specific modalities (MRI, biopsy). KER2 offers broader multi-modal integration rather than single-point metrics.

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.

5.3.5 Relevant Products and Comparative Analysis

A subset of products is particularly relevant when positioning KER2. The following table summarises representative examples and contrasts them with KER2's envisaged offering.

Table 12. Relevant products and comparative analysis with KER2

Product	Vendor	Main function	Target segment	Comparison with KER2
Owkin Data[70]	**Owkin**	Federated learning on distributed datasets to train AI models for biomarker discovery and trial optimization.	Pharma R&D and clinical researchers (oncology focus).	**Direct Competitor.** Owkin uses similar federated tech but focuses on cancer. KER2 differentiates via deep specialization in liver/metabolic disease and integration with mechanistic models.
Tempus Lens[71]	**Tempus**	Massive library of multi-modal data and AI tools for precision medicine and patient stratification.	Oncologists, Healthcare providers, Pharma R&D.	**Competitor / Complement.** Relies on centralized data. KER2's federated, vendor-neutral architecture appeals to EU consortia wishing to retain data sovereignty.
Liver-MultiScan[72]	**Perspectum**	AI-driven MRI biomarkers (cT1, PDFF) for liver health, regulatory cleared.	Clinical diagnostics, Pharma trials.	**Niche Competitor.** Focuses purely on imaging quantification. KER2 is broader, integrating MRI metrics with omics/clinical data to find composite patterns.
Syngo Carbon Intelligent Workflow[73]	**Siemens Healthineers**	Clinical decision support integrating imaging, lab, and genetic data to guide treatment pathways.	Clinicians (multi-disciplinary teams), Hospital IT.	**Indirect Competitor / Future Target.** Focuses on workflow and guidelines. KER2 focuses on exploratory clustering and novel phenotype discovery beyond current guidelines.
SOPHIA DDM[74]	**SOPHIA GENETICS**	Cloud platform unifying genomic, imaging, and clinical data for collaborative analytics.	Academic medical centers, BioPharma research.	**Potential Partner.** A generalist platform. KER2 can serve as a specialized "liver analytics module" running on SOPHIA, enhancing their offering.
Evidence[75]	**Unlearn.AI**	AI-powered solutions to reduce trial size, increase power, accelerate timelines.	Pharma trial design teams, Regulators.	**Adjacent / Partial Competitor.** Focuses on trial reduction via synthetic controls. KER2 focuses on **stratification** and understanding disease mechanisms,

[70] https://www.owkin.com/federated-research-network
[71] https://www.tempus.com/life-sciences/lens
[72] https://www.perspectum.com/our-products/livermultiscan
[73] https://www.siemens-healthineers.com/es/digital-health-solutions/intelligent-workflow
[74] https://www.sophiagenetics.com/sophia-ddm/
[75] https://www.unlearn.ai/evidence

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.

	which creates the high-quality cohorts Unlearn needs.

5.3.6 Market Dynamics and Value Network

The adoption of KER2 is influenced by specific drivers and barriers in the healthcare ecosystem.

Procurement Processes: Similar to KER1, this favours the Pharma/CRO market initially. Pharma R&D procurement is value-driven and agile, seeking immediate solutions to high trial failure rates. Hospital procurement for clinical diagnostics is complex and risk-averse.

Data Governance & Privacy: The complexity of multi-centre data sharing (GDPR) is a major barrier for centralized AI but a driver for KER2. Its federated learning architecture turns this barrier into an advantage by allowing analysis without moving patient data.

Validation & Trust: KER2 mitigates "black box" scepticism by incorporating uncertainty quantification and linking clusters to biological mechanisms, fostering trust among clinicians.

Policy & Ecosystem: The European Health Data Space (EHDS) and Virtual Human Twin initiatives are significant enablers, creating a policy environment that demands the exact harmonization and modelling capabilities KER2 provides.

Table 13. Barriers and enablers for KER2 adoption

Factor	Barrier or Enabler?	Impact on KER2 Adoption	Possible Mitigation or Leverage Strategy
Unmet Need in NASH Trials	**Enabler (Major)**	High failure rates in liver trials drive pharma to seek better stratification tools.	**Leverage:** Position KER2 as a risk-reduction tool that identifies responders and creates "cleaner" trial cohorts.
Data Privacy (GDPR)	**Barrier & Enabler**	Prevents easy data pooling but drives demand for privacy-preserving tech.	**Leverage:** Market KER2's federated architecture as the solution to GDPR constraints, enabling multi-site research.
Federated analysis challenges	**Barrier**	Operational aspects related to need for need for synchronisation and cybersecurity policy in the user organizations	**Leverage:** Expand the federated network created during the project.
"Black Box" Skepticism	**Barrier**	Clinicians wary of opaque AI clusters they cannot explain.	**Mitigate:** Emphasize KER2's explainability features (feature importance) and validation via mechanistic models.
Vendor Lock-in	**Barrier**	Hospitals/Pharma tied to closed ecosystems.	**Mitigate:** Promote KER2's vendor-neutral, containerized design that works across different hardware and platforms.
EU Strategic Initiatives	**Enabler**	Creates demand for tools that harmonize data and support digital twins.	**Leverage:** Align KER2 development with EU standards to become a reference tool for the Virtual Human Twin data layer.

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.

6 KER3 – Multiorgan Multiscale Multilevel model of liver-heart axis for MAFLD -Virtual Twin models

6.1 Functional analysis

The development of "Multiorgan Multiscale Multilevel model of liver-heart axis for MAFLD -Virtual Twin models" (KER 3) is fundamentally motivated by the need to address the profound complexity and heterogeneity of chronic metabolic diseases, particularly the challenge of personalized patient management in the context of multiorgan dysfunction.

The main scientific and market needs that drive the necessity for developing such integrated virtual twin models include:

6.1.1 Personalized Prediction and Management of Complex Disease Trajectories

The core motivation is the clinical necessity to move beyond standard care protocols designed for "average patients" toward individualized therapeutic strategies.

- Simulating Individual Pathophysiology: These models are essential to exploit patient real-world multimodal data (including patient history, imaging, histology, and omics profiles) to simulate the unique pathophysiology of the individual patient. This capacity is crucial for adapting "generic Virtual Twins" into "patient-specific Virtual Twins".
- Predicting Disease Progression and Complications: The models are developed specifically to predict the evolution of the disease, including progression into advanced forms such as cirrhosis or hepatocellular carcinoma (HCC). This includes providing predictions of Major Adverse Cardiac Events (MACE) and other cardiovascular complications.
- Assessing Treatment Efficacy: A key application is predicting the patient's response to specific therapeutic interventions, such as drug treatments (e.g., statins, SGLT2 inhibitors), minimally invasive procedures (e.g., TIPS placement), or major surgeries (e.g., liver transplantation).

6.1.2 Enhancing Mechanistic Understanding of Multi-Organ Crosstalk

From a scientific standpoint, these multiscale, multiorgan models are required to formalize and test complex mechanistic hypotheses regarding disease etiology and progression that cannot be addressed by single-scale or single-organ models alone.

- Understanding the Liver-Heart Axis: A major scientific challenge is that a significant proportion of MAFLD patients die from cardiovascular comorbidities. The multiorgan VT models address the critical need to better understand the communication mechanisms between the liver and heart. This is achieved by modelling their interaction through the bloodstream via soluble factors (growth factors, cytokines, ECM fragments) or circulating cells (macrophages, fibroblasts).
- Integrating Multi-Level Physiology: VTs provide enhanced knowledge by integrating models ranging from the molecular and intracellular scale (signal transduction, metabolism, genetic mechanisms) up to the tissue, organ (liver/heart), and whole-body circulation level. This integration closes the gap between mechanistic knowledge, often derived from animal or *in vitro* models, and clinically accessible patient data.

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.

6.1.3 Accelerating Clinical Translation via Decision Support

The integration of VTs is motivated by the market need to embed these advanced computational tools into a practical Clinical Decision Support System (CDSS).

- Delivering Actionable Knowledge: The models are geared to give responses to specific clinical questions, providing actionable knowledge to clinicians. The system facilitates the interpretation of the patient's multimodal data gathered in clinical practice.
- Risk Stratification: The VTs assist clinicians in achieving better risk stratification of high-risk patients (e.g., those at risk of fast fibrosis progression or cardiac complications) to enable timely interventions or preventive measures.
- Trustworthiness and Explainability: By combining mechanistic modeling (which implements biological consensus and hypothesized mechanisms) with AI/machine learning models, VTs strengthen the explainability and interpretability of predictions. This is necessary to leverage clinicians' trust in computational predictions before adoption in routine practice.

6.1.4 Technical Advancement and Sustainability

The complexity of the multiorgan model inherently advances the state of the art in computational infrastructure for biomedical research.

- Computational Efficiency: The VTs motivate the combination of AI and mechanistic models to optimize simulation times, making complex simulations fast enough to be compatible with clinical workflow requirements (e.g., ideally within seconds).
- Standardization and Ecosystem Contribution: The development contributes to collective efforts toward international standards (e.g., ISO/DTS 9491) and European coordination efforts, such as the Virtual Human Twin roadmap (EDITH project), ensuring that the models are interoperable, reusable, and sustainable.

6.2 Technology Analysis

6.2.1 Limitations of the current state of the art

Despite important advances in computational modelling and AI, the current state of the art still falls short of what is required to support multiorgan, patient-specific decision making in MASLD and related conditions. Recent systematic and scoping reviews of human digital twins in healthcare describe a rapidly expanding landscape of pilots and early clinical studies, mainly in cardiology, oncology and diabetes, but emphasise that most implementations remain small-scale, disease-specific and rarely multiorgan[76]. First MASLD-related digital twin interventions have been tested (e.g., a digital twin-enabled personalised nutrition programme that improved HbA1c, liver fat and fibrosis scores in people with type 2 diabetes and MASLD) but these solutions operate as data-driven metabolic twins rather than mechanistic, multiscale models of the liver–heart axis[77].

[76] https://www.nature.com/articles/s41746-024-01146-0

[77] https://www.sciencedirect.com/science/article/pii/S1530891X23005438

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.

6.2.1.1 Lack of sufficient integration and scope

Existing models lack sufficient integration and scope. Mechanistic liver models and emerging liver digital twins already capture detailed intrahepatic processes such as zonated metabolism, perfusion and regeneration after drug-induced damage, but they are typically designed as single-organ systems[78]. In parallel, high-fidelity heart digital twins have reached a relatively advanced stage: platforms such as Dassault Systèmes' *Living Heart* are commercially available and used worldwide for cardiovascular device design, testing and in silico trials[79]. Similarly, industry prototypes such as Siemens Healthineers' *liver digital twin* are being developed for surgical planning in liver cancer and fatty liver disease[80].

However, these developments remain largely organ-specific. Multiorgan models that explicitly represent the liver–heart axis and the circulation of soluble mediators and cells, instantiated as patient-level Virtual Twins, are still rare and fragmentary. Recent multi-organ imaging work in large cohorts, such as MRI-based studies of the heart–brain–liver axis in cardiometabolic disease, underlines the biological importance of interorgan crosstalk but does not yet provide integrated, mechanistic Virtual Twins that couple liver and heart dynamics[81]. Even within the liver, many submodels at molecular, cellular, tissue and organ level are still developed in isolation and not fully integrated into a single predictive twin that can be interrogated in near real time.

From a computational perspective, purely mechanistic models remain heavy and slow, often with simulation times incompatible with clinical workflow. In cardiology, large-scale "populations" of digital heart twins generated from imaging and ECG data have demonstrated that surrogate models and Gaussian-process emulators can reduce simulation times from hours to minutes, making thousands of personalised heart models feasible[82]. Yet similar acceleration strategies and hybrid mechanistic–AI frameworks are not yet standardised for clinically oriented, multiorgan MASLD Virtual Twins that must incorporate both hepatic and cardiac physiology and their interactions under different interventions.

6.2.1.2 Data limitations significantly constrain personalisation and validation

MASLD is now recognised as a multi-organ condition with strong links to cardiovascular, renal and other extrahepatic diseases. Recent reviews detail the complex inter-organ crosstalk and extrahepatic complications associated with MASLD, but implicitly highlight the paucity of longitudinal, deeply phenotyped cohorts that capture these interactions in a way suitable for mechanistic modelling[83]. Large population studies with multi-organ imaging (for example, UK Biobank-based analyses of heart–brain–liver relationships) begin to characterise organ-to-organ associations at scale, yet they typically lack the multimodal combination of histology, omics, ECG and intervention outcomes needed to parameterise and externally validate Virtual Twins of the liver–heart axis[84]. For mechanistic models in particular, the scarcity of patient-specific inputs such

78 https://www.researchgate.net/publication/374270968_A_liver_digital_twin_for_in_silico_testing_of_cellular_and_inter-cellular_mechanisms_in_regeneration_after_drug-induced_damage

79 https://www.3ds.com/3dexperiencelab/portfolio/living-heart

80 https://www.siemens-healthineers.com/perspectives/modeling-the-human-liver

81 https://pubmed.ncbi.nlm.nih.gov/40403110/

82 https://www.electropages.com/blog/2025/07/digital-twins-human-heart-help-researchers-diagnose-heart-disease

83 https://pubmed.ncbi.nlm.nih.gov/36904122/

84 https://pubmed.ncbi.nlm.nih.gov/40403110/

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.

as serial biopsy-derived omics or detailed haemodynamic measurements remains a major bottleneck for robust personalisation and out-of-sample validation.

6.2.1.3 Translation into clinical practice is hampered by usability and trust issues

A recent scoping review of human digital twins in healthcare highlights persistent barriers around interoperability, explainability, regulatory acceptance and integration with clinical information systems, noting that most implementations remain outside everyday care pathways or are confined to research settings[85]. In metabolic disease management, both academic trials and commercial solutions show that digital-twin-based care can be deployed: the above-mentioned digital twin-enabled nutrition programme for type 2 diabetes and MASLD is one example of an intervention embedded in care pathways[86]. Likewise, companies such as Twin Health offer app-mediated "AI digital twin" services that provide continuous metabolic monitoring and behavioural guidance using wearable and lab data[87].

However, these solutions operate mainly as data-driven, system-level or single-organ twins, rather than as deeply integrated, multiscale liver–heart models capable of simulating invasive interventions (e.g. TIPS, liver transplantation) and explaining mechanistic drivers of predicted outcomes. They are also generally not embedded into hospital IT infrastructures as transparent, query-able components of Clinical Decision Support Systems (CDSS) for chronic liver disease, despite their potentially critical role in clinical decision and their mandatory classification as medical devices (SaMD).

6.2.2 ARTEMIS results and technical features beyond the state of the art

Current limitations motivate the development of KER3 within ARTEMIS: an integrated multiorgan, multiscale liver–heart Virtual Twin for MASLD that (i) consolidates and extends existing submodels across scales and organs, (ii) is parameterised and tested on a large, multimodal, multicentric cohort, and (iii) is designed from the outset for explainable, efficient use in a CDSS. This design is aligned with the directions set out in the European Virtual Human Twin (VHT) roadmap, which calls for interoperable, disease-specific virtual twins that can be integrated into a broader European ecosystem for Virtual Human Twins in health and care[88]. In combination with the processing pipelines (KER1), multimodal clustering and pattern identification (KER2) and the CDSS framework (KER5), KER3 is conceived to overcome the current fragmentation of models and data and to provide a credible pathway from advanced computational modelling to clinically usable, trustworthy tools.

[85] https://www.nature.com/articles/s41746-025-01910-w
[86] https://www.sciencedirect.com/science/article/pii/S1530891X23005438
[87] https://usa.twinhealth.com/for-health-plans
[88] https://www.edith-csa.eu/roadmap/

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 14..Main results clustered under KER3

ID	WP Task(s)	Main partners contributing	Result / component	Description and functionality	Technical features beyond the state of the art	Exploitation potential and link to KER3
R1	T5.1	MUV, BU	**ML submodels and MRI progression pipeline for liver disease course prediction**	Pre-processing and feature-extraction pipeline for liver MRI, combining segmentation, vesselness analysis and registration, applied to sequential scans to characterise quantitative trajectories of tissue changes. Built on these features, supervised ML models (including explainable methods) are used to predict progression of steatosis and fibrosis and identify key clinical and imaging predictors.	Moves beyond static, cross-sectional risk scores by explicitly modelling *paths* through tissue types in longitudinal MRI and linking them to clinical features. Incorporates explainable AI (Local Interpretable Model-Agnostic Explanations - LIME) and is designed to evolve into distributed / federated ML on multimodal, longitudinal data, addressing the SOTA gap in robust, multi-centre MAFLD progression models.	Provides the data-driven disease course prediction layer of KER3, complementing mechanistic VT components with ML-based risk scores for steatosis/fibrosis progression. It supports risk stratification in use cases (e.g. fast progressors) and can be repackaged as a standalone service for liver MRI analytics or embedded in the CDSS as a prognosis module.
R2	T5.2	ULEI, BU	**Physics-informed deep-learning Reduced-Order Models (ROMs) for mechanistic VT submodels**	Systematic evaluation of Deep Learning (DL) based approaches to effectively approximate the analytical solution of complex mathematical equations (e.g., ODEs and PDEs) used in mechanistic models (intracellular, cellular, tissue scales). These ROMs function as computationally efficient approximations of the high-	Utilizes advanced techniques such as Physics-Informed Neural Networks (PINNs) and Fourier Neural Operator (FNO) approaches to solve differential equations. By leveraging physics-constrained methods (incorporating initial and boundary physical conditions) alongside data-constrained	This technological advancement is anticipated to lead to drastically improved model simulation performance, which is required to achieve runtimes compatible with usage in clinical practice. The ROMs enable rapid patient-specific predictions, facilitating the broad utilization of the

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.

				dimensional dynamics of the full mechanistic submodels.	approaches, the ROMs are constructed as faithful, yet significantly faster, representations of the complex biological processes.	Virtual Twin models within the Clinical Decision Support System (CDSS) prototype.
R3	T5.3	DKFZ, ALU-FR, INRIA, ULEI, UKHD	**Intracellular signalling and metabolism models for liver–heart communication**	Dynamic ODE-based pathway models for key mediators (HGF for metabolism, TGF-β for tissue remodelling, IL-6 for inflammation), calibrated using high-throughput proteomic data. A workflow integrates proteomics, imaging and ML (e.g. MRMR feature selection) to identify molecular panels associated with steatosis and fibrosis progression and to quantify drug effects (e.g. common painkillers influencing hepcidin via IL-6 signalling).	Goes beyond SOTA by linking proteomic signatures and intracellular signalling to secreted factors and downstream tissue mechanics, tailored to disease subgroups. The subgroup-specific calibration of ODE models to proteomic panels and explicit demonstration of drug-induced changes in iron metabolism provide a mechanistic bridge from omics to functional outputs relevant for the liver–heart axis.	Supplies the molecular-scale layer of KER3, generating mechanistic inputs (e.g. time-courses of cytokines, growth factors and ECM-related outputs) for tissue, hemodynamic and organ-level VT components (T5.4, T5.5). It underpins future exploitation in drug response simulation, identification of molecular targets, and personalised scenario testing (e.g. impact of specific medications on liver and heart outcomes).
R4	T5.4	INRIA-Drasdo, ULEI, ALU-FR, DKFZ, ICAN, UKHD, VHIR	**Spatio-temporal liver and myocardium tissue Virtual Twins with generalized**	Development of spatio-temporal submodels for liver and heart tissue, including a generalized collagen network model that simulates secretion, diffusion, adhesion, interlinking and digestion of collagen fibres. These models reproduce biomechanical responses observed experimentally	Provides a unified collagen network and ECM remodelling framework applicable to both liver and heart fibrosis, validated against biomechanical experiments—something not available in existing organ-specific models. The multi-level tissue	Forms the tissue-scale backbone of KER3, enabling simulation of fibrosis progression and potential reversibility under different scenarios (drugs, TIPS, transplantation). It can be exploited in further research on fibrosis mechanisms and

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.

			collagen network	and extend the Liver VT with additional components (lipid droplets, transaminases, IL-6, HGF, TNFα) to mimic progression from steatosis to fibrosis in both liver and myocardium.	VT spans molecular to micro-architectural scales and directly addresses the gap between histology and clinical imaging/biomarkers identified in the SOTA.	has clear commercial potential as a specialised VT module for pharma (anti-fibrotic therapies) and device companies interested in tissue-level mechanical effects.
R5	T5.5	INRIA-Vignon, ULEI, JUH, AP-HP	**Multiscale hemodynamics and transport models (HeTran-TL/OL/BL) for liver–heart coupling and interventions**	Suite of hemodynamic and transport models at tissue (HeTran-TL), organ (HeTran-OL) and whole-body level (HeTran-BL). Includes a closed-loop lumped-parameter circulation model with explicit liver and heart representation using the Single Fibre Model (SFM), preliminary CFD-based TIPS simulations, and a reduced-order patient-specific TIPS shunt model. These models quantify blood flow, pressure and transport of soluble factors (e.g. ammonia, signalling molecules) and assess the impact of interventions such as TIPS and transplantation.	Represents one of the first integrated liver–heart circulation models tailored to portal hypertension and TIPS, including investigation of how TIPS modifies natural porto-systemic collaterals and the fraction of shunt flow originating from ammonia-rich mesenteric veins. The combination of CFD-derived insights with lumped-parameter and reduced-order models goes beyond existing single-organ or generic cardiovascular models.	Provides the organ- and whole-body-scale coupling layer of KER3, critical for simulating post-TIPS heart failure risk, portal pressure gradients and systemic consequences of liver interventions (UC2–UC3). This can evolve into a clinical decision-support tool for TIPS planning (graft diameter/site selection) and a reusable hemodynamic engine for other VT-based products.
R6	T6.1	INRIA-Drasdo, DKFZ, INRIA-Vignon, ALU-FR, ULEI	**Multilevel orchestration framework and VT templates**	Definition and implementation of the orchestration layer that couples intracellular, tissue, organ and whole-body submodels into coherent Virtual Twins tailored to specific use cases (UC1–UC4).	Goes beyond SOTA by providing an explicit, reusable orchestration architecture for multi-organ, multiscale VTs, rather than ad-hoc model combinations. It combines	Delivers the integrated KER3 Virtual Twin prototypes for each clinical use case, ready to be linked with KER1 (pipelines), KER2 (patient clusters) and KER5 (CDSS).

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.

			for liver–heart use cases	Includes schematics and implementations for couplings such as hepatic encephalopathy (0D hemodynamics + transport + ammonia detoxification models), and a first disease-progression tissue model from fatty liver to fibrosis integrated into the VT framework. AI-driven functional mappings are being developed to connect complex pathway inputs/outputs efficiently.	mechanistic couplings (via circulating factors/cells) with AI-based connectors for computational efficiency, aligned with realistic clinical questions and constraints defined in the use cases.	The orchestration framework can be further exploited as a generic VT integration layer for other diseases and as a methodological asset in the emerging Virtual Human Twin ecosystem.
R7	T6.4	INRIA, DKFZ, ALU-FR, ULEI	**FAIR-compliant modelling platforms (CompuTiX and LumpedFlux) for 0D circulation and micro-architectural tissue models**	Development of LumpedFlux, a modular C++ software package for 0D circulation models that can be coupled with 3D CFD and other simulators, and CompuTiX[89], a fourth-generation micro-architectural modelling framework for cell-based and tissue-level simulations. Both are designed according to FAIR principles, with code architectures that facilitate coupling, extension and community use, and are foreseen as core components in the Virtual Human Twin (VHT) ecosystem.	Addresses a structural SOTA gap by providing open, modular and standard-ready platforms for multiscale VT development. The combination of mathematically robust 0D circulation (LumpedFlux) and flexible, 3D-capable tissue modelling (CompuTiX) under FAIR and VHT-aligned requirements is not available in existing proprietary or siloed solutions.	Ensures the sustainability and interoperability of KER3, enabling its models to be maintained, extended and integrated with external VT initiatives (e.g. EDITH VHT roadmap) and future ISO/DTS standards. Both tools have exploitation potential as community platforms and as embeddable engines for third-party VT applications in academia, hospitals and industry.

[89] https://team.inria.fr/simbiotx/software-2/computix/

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.

6.3 Market analysis

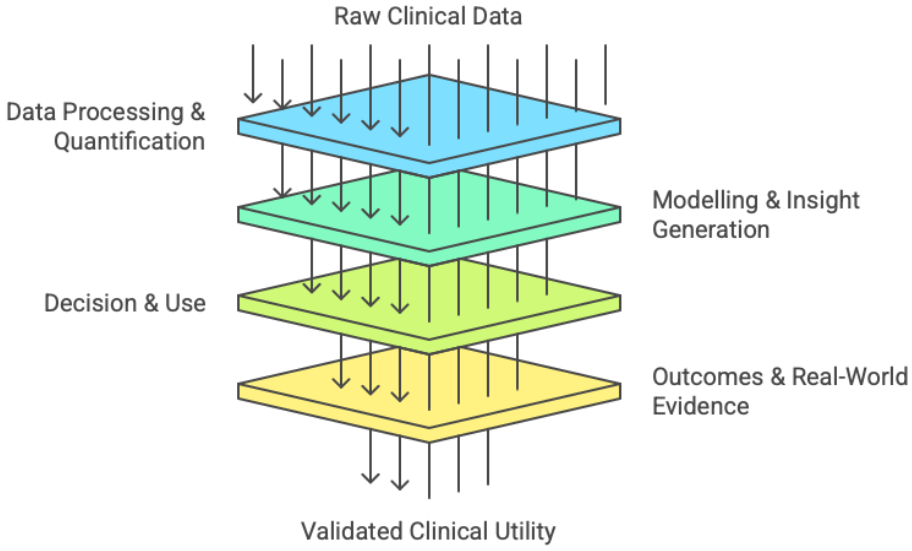
6.3.1 Value chain overview

The value chain surrounding KER3's multi-organ virtual twin technology begins upstream with **Data Processing & Quantification**. The primary actors at this stage are quantitative imaging labs, AI-based histology providers (like the KER1 solution), and hospital IT systems. Their core activity is transforming raw clinical data (imaging, labs, omics) into standardized, structured features. This process results in an output of model-ready parameters. However, these metrics remain descriptive and isolated, lacking the dynamic, systemic interconnectivity required to predict complex disease trajectories.

The second link in the value chain is **Modelling & Insight Generation**. This is the layer where KER3 creates its primary value. Its core activity involves ingesting harmonized multi-modal data to initialize biophysical and AI-driven "virtual twins" of the liver-heart axis. Unlike simple risk scores, KER3 simulates time-dependent disease progression (e.g., fibrosis evolution) and organ crosstalk (hemodynamic interactions). This process results in an output of patient-specific predictive trajectories and "what-if" scenario simulations. This output is characterized by its ability to forecast outcomes under different interventions, bridging the gap between static data and dynamic biological understanding.

The third link in the value chain is **Decision & Use**. The primary actors at this stage are Pharma R&D teams, Clinicians (hepatologists, cardiologists), and Digital Health providers. Their core activity involves applying KER3's simulations to high-stakes decisions: optimizing clinical trial designs (in-silico trials), selecting personalized therapies, or guiding patient lifestyle changes. This process results in an output of de-risked R&D pipelines and personalized care plans.

The final link in the value chain is **Outcomes & Real-World Evidence**. The primary actors here are Patients, Payers and Regulators. Their core activity is monitoring the actual health improvements and cost efficiencies resulting from the twin-guided decisions. This process results in an output of validated clinical utility and health-economic data, which provides the necessary feedback loop to refine the upstream KER3 models and support reimbursement.



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.

6.3.2 Potential application fields for the KER

Given ARTEMIS' focus on the Liver-Heart axis in MAFLD, the most relevant application fields are those where predictive modelling of multi-organ complications can reduce trial failure rates, personalize complex chronic care, or engage patients. Below, we summarise these fields and narrow down to the segments with the highest potential.

6.3.2.1 Pharma R&D: In-silico trials for NASH and cardio-metabolic drugs

The most attractive application field for KER3 is the simulation of drug effects for pharmaceutical R&D, specifically for NASH and related cardio-metabolic conditions. The global NASH therapeutics market is projected to grow from ~$7.7 billion in 2024 to over $33 billion by 2030 (CAGR ~28%)[90]. However, these trials face high failure rates and difficulty in identifying responders.

Regulators and industry are increasingly embracing in-silico approaches to supplement physical trials, reducing costs and ethical burdens. KER3's ability to act as a virtual patient simulator allows sponsors to 1) identify biomarkers of response, 2) optimize trial design by selecting patients least likely to respond to placebo, and 3) predict long-term cardiovascular outcomes beyond standard trial duration. This aligns with the broader Digital Twins in Healthcare market, which is projected to reach ~$60 billion by 2030[91]. For KER3, the immediate opportunity is to provide simulation services to de-risk Phase II/III decisions for major players like Novo Nordisk or Eli Lilly.

6.3.2.2 Clinical Decision Support (CDSS) for MAFLD and comorbidities

A second core field is the deployment of KER3 as a predictive engine within Clinical Decision Support Systems (CDSS) in hospitals and liver clinics. The global CDSS market is estimated at $4-10 billion by 2030[92] , driven by the need to manage the rising burden of chronic metabolic diseases.

Clinicians lack tools to accurately predict which MAFLD patients will develop life-threatening cardiovascular complications, which are the leading cause of death in this population. KER3 addresses this by simulating liver-heart interactions to forecast disease progression (e.g., fibrosis, heart failure). This enables "active surveillance" and personalized treatment planning (e.g., predicting risks of TIPS procedures). While regulatory barriers (MDR certification) exist, the integration of such predictive twins into workflows is the future of precision hepatology.

6.3.2.3 Patient engagement and digital therapeutics

A third field involves patient-facing digital health tools. The market for digital diabetes management alone is projected to reach $32 billion by 2030[93]. Trends show a shift toward "participatory medicine," where patients visualize their own "digital twin" to understand risks.

KER3 can be embedded into digital therapeutics apps to serve as an educational and coaching tool. By visualizing how lifestyle changes (diet, exercise) impact their specific liver and heart trajectory, patients are more motivated to adhere to treatment. Success stories like Twin Health,

[90] https://www.grandviewresearch.com/industry-analysis/non-alcoholic-steatohepatitis-nash-treatment-market-report

[91] https://www.prnewswire.com/news-releases/digital-twins-in-healthcare-market-worth-us59-94-billion-by-2030-with-68-0-cagr--marketsandmarkets-302445817.html

[92] https://www.grandviewresearch.com/industry-analysis/clinical-decision-support-system-market

[93] https://www.mordorintelligence.com/industry-reports/digital-diabetes-management-market

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.

which raised >$140M for metabolic digital twins[94], validate the commercial appetite for this B2C or B2B2C model.

6.3.2.4 Multi-organ research and translational medicine

The fourth field is the academic and translational research sector. Initiatives like the Virtual Human Twin are creating an ecosystem for multi-organ models. Research institutions use KER3 to contrast hypotheses about the liver-heart axis (e.g., how hepatic inflammation drives cardiac dysfunction). While not a direct high-revenue market, it is critical for scientific validation, standardization, and maintaining KER3's position at the state-of-the-art.

6.3.2.5 Integrated chronic care platforms (Diabetes/Obesity)

Finally, KER3 has potential in the broader integrated care market for metabolic syndrome. Clinical platforms managing diabetes and obesity (e.g., Virta Health, Onduo) could integrate KER3 to enhance their risk prediction capabilities. As GLP-1 agonists and SGLT2 inhibitors become standard for multi-morbid patients, a model that predicts combined liver-heart benefits offers significant value to endocrinologists and insurers.

6.3.3 Prioritisation and narrowing of target markets

Considering technical fit, unmet need, purchasing power, and regulatory barriers, the segments with the highest potential for commercialisation of KER3 are:

1. **Pharma R&D (In-silico trials):** Highest willingness to pay, aligns with urgent need to de-risk high-stakes NASH trials, lower regulatory barrier than clinical devices.
2. **Clinical Decision Support (CDSS):** When accessible, once obtained the MDR certification, this segment is characterised by a strong unmet clinical need (mortality prediction), high volume of patients andlong-term recurring value.
3. **Digital Health / Patient Apps:** Rapidly growing market, opportunity for licensing the engine to existing metabolic platforms, drives patient behavioral change.

Table 15. Analysis of market segments for KER3

Segment	Main unmet need	KER3 offer	Market characteristics	Priority (by 2027)
1. Pharma R&D (In-silico trials)	High failure rates in NASH/MASH trials, need to identify responders and de-risk high-stakes R&D investments.	Virtual patient simulation, in-silico trial cohorts, identification of mechanistic response biomarkers.	Global NASH market >$33B (2030), High CAGR (~28%).	High - Primary
2. Clinical Decision Support (CDSS)	Inability to accurately predict life-threatening cardiovascular complications in MAFLD, lack of	Personalized prognosis of liver-heart interactions, dynamic simulation of disease trajectories (fibrosis, heart failure).	Global CDSS market ~$10B (2030), Strong push for precision medicine.	High - Strategic

[94] https://www.prnewswire.com/news-releases/twin-health-announces-53m-investment-as-ai-digital-twin-unlocks-sustainable-diabetes-and-obesity-outcomes-while-reducing-medications-302535848.html

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.

	guidance for complex cases.			
3. Digital Health / Patient Apps	Low patient adherence, difficulty engaging patients in long-term lifestyle changes for chronic metabolic disease.	Visual "digital twin" for coaching, simulation of lifestyle impacts (diet/exercise) on personal health trajectory.	Digital diabetes market ~$32B (2030), Validated by recent investments.	Medium - Partnership
4. Research & Translational	Limited understanding of the mechanistic link between hepatic inflammation and cardiac dysfunction, need for biomarkers.	Multi-scale modeling of Liver-Heart axis mechanisms, hypothesis testing environment for academic consortia.	Driven by public funding (Horizon Europe), "Blue Ocean" for scientific discovery.	Strategic - Foundation
5. Integrated Chronic Care	Fragmented care for multi-morbid patients (diabetes + obesity + NAFLD), lack of holistic risk stratification.	Combined risk stratification for metabolic syndrome, prediction of dual-organ benefits from therapies (e.g., GLP-1).	High volume, crowded market requires integration with major platforms.	Low Short-term

6.3.4 Identification of key players in the value chain

The following table summarises key archetypes and example organisations relevant for KER3.

Table 16. Key players in the KER3 value chain

Role in value chain	Type of organisation	Example players	Relevance to KER3
End customers - Pharma & Biotech	Companies developing NASH, diabetes, or cardio-metabolic drugs.	**Novo Nordisk**, **Eli Lilly**, **Madrigal**, **Gilead Sciences**, **Pfizer**.	Primary buyers for in-silico trial services to test drug efficacy and safety in virtual populations.
End customers - Hospitals & Providers	Academic Medical Centres, Hepatology/Cardiology clinics.	**Vall d'Hebron**, **AP-HP Paris**, **EASL** affiliated centers.	Adopters of KER3 for clinical decision support and complex patient management.
Providers & Partners - Health IT	EHR vendors, Imaging PACS, Cloud infrastructure.	**Epic**, **Oracle Cerner**, **GE HealthCare**, **Philips**, **AWS**, **Azure**.	Critical channels for integration into clinical workflow, hosting of computationally intensive simulations.
Providers & Partners - Data/AI	Data integration, Modelling platforms, AI analytics.	**Medexprim**, **Ansys**, **OpenVHP**.	Enable data ingestion (HL7/FHIR) and provide complementary modelling toolkits[95].

[95] https://www.ansys.com/events/dm-digitaltwins

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.

Direct Competitors - Virtual Twins	Virtual Human Twin platforms and organ-specific models.	**Dassault Systèmes** (Living Heart/Liver), **Nova Discovery** (Jinkō).	Major competitors offering sophisticated physics-based simulation platforms for pharma and devices.
Substitute Solutions - Digital Health	AI-driven metabolic twins and coaching apps.	**Twin Health**[96], **Virta Health**, **Simulations Plus** (NAFLDsym).	Provide alternative metabolic management or pharma simulation without the specific liver-heart coupling.
Regulators & Ecosystem	Regulatory agencies, Standard bodies, Consortia.	**EU Commission, MDCG, EMA**, **FDA**, **EDITH**, **VPH Institute**.	Gatekeepers for clinical/trial validity, setters of standards (e.g., ASME V&V 40) for computational models.

This mapping highlights that KER3 operates in a high-value intersection between **Pharma R&D** (competing with specialized QSP models like NAFLDsym) and **Clinical Care** (competing with or complementing major Digital Twin players like Dassault).

6.3.5 Relevant products and comparative analysis

A subset of products is particularly relevant when positioning KER3. The following table summarises representative examples and contrasts them with KER3's envisaged offering.

Table 17. Relevant products and comparative analysis with KER3

Product	Vendor	Main function	Target segment	Comparison with KER3
Living Heart / Living Liver[97]	**Dassault Systèmes**	High-fidelity 3D physics simulation of organ function for surgical planning and device testing.	Device makers, Surgeons, Researchers.	**High-End Competitor.** Focuses on detailed anatomy/surgery. KER3 differentiates by focusing on **chronic metabolic progression** and the **systemic Liver-Heart axis** rather than acute surgical intervention.
Whole Body Digital Twin[98]	**Twin Health**	AI-driven metabolic twin using sensors (CGM, wearables) to reverse diabetes.	Patients, Payers, Employers.	**AI/Data Competitor.** Uses AI/ML for lifestyle coaching[99]. KER3 adds value via **mechanistic biophysical modelling** (predicting organ damage/hemodynamics) rather than purely data-driven correlation.
FFRct[100]	**Heart-Flow**	Non-invasive simulation of coronary blood	Cardiologists,	**Clinical Benchmark.** A regulated, reimbursed diagnostic. KER3 targets a broader **multi-organ systemic**

[96] https://www.prnewswire.com/news-releases/twin-health-announces-53m-investment-as-ai-digital-twin-unlocks-sustainable-diabetes-and-obesity-outcomes-while-reducing-medications-302535848.html

[97] https://www.3ds.com/3dexperiencelab/portfolio/living-heart

[98] https://usa.twinhealth.com

[99] https://pubmed.ncbi.nlm.nih.gov/37778441/

[100] https://www.heartflow.com/heartflow-one/ffrct-analysis/

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 from CT to assess CAD.	Radiologists.	**risk** (liver-heart) rather than a single vessel functional metric.
NAFLDsym[101]	**Simulations Plus**	QSP software simulating liver fat, inflammation, and fibrosis pathways for drug testing.	Pharma R&D (DMPK/Clin Pharm).	**Pharma Competitor.** The standard for liver drug modelling. KER3 differentiates by integrating the **cardiovascular component**, critical for assessing safety and comorbidities in NASH.
Jinkō[102]	**Nova Discovery**	In-silico clinical trial simulation platform using virtual populations.	Pharma Clinical Development.	**Platform Competitor.** Offers trial simulation as a service. KER3 could be a **complementary module** (Liver-Heart engine) within such broader platforms rather than a direct platform rival.

Overall, KER3's competitive angle is its unique multi-organ scope (bridging Hepatology and Cardiology), whereas most competitors are siloed in one organ (HeartFlow, NAFLDsym) or focus purely on metabolic data without organ-level physics (Twin Health).

6.3.6 Market dynamics and value network

The adoption of KER3 is governed by several key market dynamics, barriers, and enablers.

Data & Interoperability: KER3 is data-hungry, requiring multi-modal inputs. The "siloed" nature of hospital data is a barrier. However, the rise of interoperability standards (HL7 FHIR) and the EHDS acts as a massive enabler, facilitating the necessary data flow.

Validation & Trust: A major barrier is the "black box" scepticism regarding AI/models. Regulators and clinicians require clear evidence and rigorous validation. KER3's alignment with ASME V&V 40 standards and participation in EU initiatives like EDITH are key strategies to build credibility.

Economic Incentives: Hospital adoption depends on reimbursement. The example of HeartFlow obtaining payer coverage after proving cost-effectiveness (reduced angiograms) is instructive. KER3 must demonstrate that preventing liver/cardiac complications saves money to unlock value-based care payments.

Table 18. Barriers and enablers for KER3 adoption

Factor	Barrier or Enabler?	Impact on KER3 Adoption	Possible Mitigation or Leverage Strategy
Regulatory Acceptance	**Barrier & Enabler**	Regulatory uncertainty for "Virtual Twins" is high, but agencies (FDA/EMA) are opening pathways for in-silico evidence.	**Leverage:** Align with EMA's Model-Informed Drug Development (MIDD) and use in-silico trials as the initial entry point before clinical diagnostic use.
Workflow Integration	**Barrier**	Clinicians will not use tools that require manual data entry or exist outside the EHR.	**Mitigate:** Develop "invisible" integration where KER3 runs in the background and pushes results directly to the EHR dashboard.

[101] https://www.simulations-plus.com/software/nafldsym/

[102] https://www.novainsilico.ai/Jinko/

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.

Evidence & Validation	**Barrier**	Skepticism regarding the accuracy of predicting long-term outcomes (e.g., 5-year fibrosis risk).	**Mitigate:** Execute robust retrospective and prospective validation studies (as planned in ARTEMIS) and publish in high-impact journals.
Ecosystem & Standards	**Enabler**	The **EDITH** and **VPH** initiatives provide a framework for standards, reducing fragmentation.	**Leverage:** Position KER3 as a reference implementation of the European Virtual Human Twin, benefiting from non-dilutive funding and network effects.

7 DATA SUSTAINABILITY STRATEGY

7.1 Datasets processed in ARTEMIS

7.1.1 Summary of the ARTEMIS cohort

The ARTEMIS cohort aggregates anonymised, multimodal, longitudinal patient data provided by the consortium's clinical centres (also named as Data Holders), to support the development and validation of VHT models for the liver–heart axis. The cohort comprises four clinical use cases reflecting the spectrum of MASLD-related disease progression from early metabolic dysfunction to advanced cirrhosis, cardiovascular complications and hepatocellular carcinoma. The cohort is aligned with the objective of enabling multicentric model development, using mechanistic and AI-based approaches.

Use Case 1 (UC1) - Leading Clinicians: Dr JM. Pericas (VHIR) and Dr R. Pais (ICAN), Leading Modeller: D. Drasdo (INRIA). UC1 focuses on fibrosis progression in MASLD. It includes adult patients diagnosed with MASLD through clinical, biochemical and radiological indicators, with a minimum 5-year follow-up. This longitudinal depth supports the study of disease trajectories and personalised modelling of fibrosis risk.

Use Case 2 (UC2) - Leading Clinicians: Pr N. Frey and Dr F. Leuschner (UKHD), Leading Modeller: U. Klingmüller (DKFZ). UC2 focuses on cardiac fibrosis associated with MASLD. It includes adult patients with MASLD and cardiovascular disease, characterised by radiological or histological evidence of cardiac fibrosis and at least 1 year of follow-up. This subgroup provides critical insights for modelling cross-organ mechanisms linking liver disease to heart failure.

Use Case 3 (UC3) - Leading Clinician: Dr. C. Ripoll (JUH), Leading Modeller: I. Vignon-Clementel (INRIA). UC3 focuses on patients with cirrhosis and portal hypertension who undergo TIPS placement or liver transplantation. Their comprehensive clinical profiles enable mechanistic modelling of haemodynamic changes and cardiovascular decompensation following major interventions.

Use Case 4 (UC4) Leading Clinicians: Dr. N. Lanthier and Pr. Y. Borbath (CUSL), Leading Modeller: S. Höhme (ULEI). UC4 focuses on cardiac complications following HCC therapy. It includes adults with HCC of any aetiology receiving treatment (surgical, interventional radiology or systemic therapy) with at least 6 months of follow-up.

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.

For each use case, a dedicated clinical study protocol has defined the inclusion/exclusion criteria, variables, data types and timelines. Clinical partners jointly agreed on a **unified list of variables** nto ensure harmonised and interoperable data (as reported in WP2). Data include clinical data, radiological imaging and when available, digital pathology and omics.

Each data holder is responsible for deidentifying their datasets before entering them into the ARTEMIS infrastructure, fully preserving patient privacy. The majority of data holders considers their deidentification process to be anonymisation meaning that the resulting dataset does not contain personal data subject to the GDPR. Some data holders, such as AP-HP consider their datasets to be pseudonymised, consequently they must inform patients about all data processes done with their data.

7.1.2 Overview of the hybrid infrastructure

The ARTEMIS cohort (the aggregated ARTEMIS dataset) is accessible to all project partners registered in the ARTEMIS project platform or data infrastructure: an environment for model development and evaluation, described in detail in *Deliverable D2.2 "Technical Description of the Data Infrastructure"*. The platform is designed as a hybrid federated/central environment enabling secure data storage, processing and exploration across all project sites. Its architecture combines **local federated nodes** deployed at clinical centres with a certified **central cloud** environment, providing flexibility for sites with different technical and regulatory constraints.

Dataset storage can be either **federated** (locally at each clinical site) or **central** (on the cloud site created for this project), as selected by the Data Holder. In both cases, access to any data is mediated exclusively through the ARTEMIS platform hosted in the cloud environment, ensuring uniform authentication and security measures. The cloud environment is a certified secure system operated under Open Telekom infrastructure and managed by MEDEX-BCP, with restricted access, ensuring that all actions are logged and traceable. The hybrid environment supports authorised user login, data analytics and model execution, while providing scalability and flexibility to manage the multimodal ARTEMIS cohort and its associated modelling workflows.

Federated learning workflows are implemented using the Flower framework, to enable AI model training and evaluation directly on local datasets.

7.1.3 Legal and governance basis

The governance of the ARTEMIS cohort is grounded in the project **Grant Agreement (GA)**, **Consortium Agreement (CA)**, and **Data Sharing Agreement (DSA)** signed by all partners. Complementing these contractual documents, the deliverables *D2.1 "Initial DMP"* and *D2.3 "Updated DMP"* define the structure of the cohort and the rules for data management and access policies, while *D2.2* describes the technical implementation of the hybrid infrastructure and the mechanisms for secure and federated data use.

The **DSA** includes specific articles governing post-project data retention:

- Article 7 (Storage limitation) requires partners to retain or process health data only for the duration necessary to fulfil the agreed project purposes. At project end, datasets must be treated in accordance with the final DMP (*D2.4*). Unless otherwise stated, datasets exchanged between partners must be deleted.
- Article 9 (Transfers of data) specifies that, unless otherwise agreed in the final DMP or bilateral agreements, partners must destroy health data they have received from other partners within 60 days after project termination and confirm destruction in writing.

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 updated DMP states that **data re-use policies and conditions for access will be further developed in the final DMP** (*D2.4*).

Mechanisms for access control and audit tracks are already operational in the hybrid infrastructure. The technical safeguards, together with the partner legal commitments offer a strong baseline upon which a long-term sustainability strategy can be developed.

7.1.4 Map of data holders

7.1.4.1 Considerations on data governance and access, in the context of use by external users

The metadata for the ARTEMIS shared datasets (stored both centrally and in federated nodes) will be publicly available via the ARTEMIS website. This site will also inform of the access request process, the contact email for reception of access requests and the terms and conditions of data use.

Interested party can submit a request for access to one or several preidentified datasets. Preliminary access-request process (to be refined in next period) includes:

- Acceptance of Terms and Conditions of use of ARTEMIS datasets. A terms and conditions policy for the datasets will be annexed in D9.4.
- Data Access request form, including user identity, dataset(s) requested, intended use, and security/ethics documentation. A form template will be annexed in D9.4.

Authorised users and permitted uses relate to **research projects only**. An extension of the authorised uses to follow the principles of the European Health Data Space (which permits data reuse for research, innovation, training, AI development/testing, regulatory decision-making, and public-interest objectives) will be planned in the final sustainability plan if supported by the Data Holders.

Procedures for access request evaluation: An Access Evaluation Committee will be formed with a minimum of 2 members: 1 ARTEMIS Coordinator representative (MAT) and 1 platform manager representative (MEDEX-BCP), with voluntary participation of 1 representative per involved Data Holder site. The evaluation criteria will consider scientific relevance of the intended use/project, societal impact of the pursued results, soundness of the implementation plan and experience of the applicant(s).

Embargo period: External access will begin only after an embargo to allow partners to finalise analyses and publications. The duration of the embargo period will be determined in D9.4 and will not be superior to 18 months.

7.1.4.2 Map of Data Holders

For each Data Holder, Table 19 shows their storage type, use cases covered and preliminary position regarding data access after project end. This map will be revised in D9.4 according to the final achievements of the project.

Table 19. Map of Data Holders

HULAFE
Storage type: Central & Federated

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.

Use cases: UC1, UC3 TIPS, UC3 LT, UC4
Access for partners: YES
External access: Restricted
VHIR
Storage type: Federated
Use cases: UC1, UC3 TIPS, UC4
Access for partners: YES
External access: Restricted
JUH
Storage type: Federated
Use cases: UC3 TIPS, UC4
Access for partners: YES
External access: Restricted
UKHD
Storage type: Federated or central
Use cases: UC1 & UC2
Access for partners: Restricted
External access: Restricted
ULEI
Storage type: Central
Use cases: UC1, UC4, Controls
Access for partners: YES
External access: Restricted
CHA
Storage type: (to be confirmed)
Use cases: UC1, UC4
Access for partners: YES
External access: Restricted
ICAN
Storage type: Federated
Use cases: UC1
Access for partners: YES
External access: Restricted
APHP Pitié Salpêtrière
Storage type: Federated
Use cases: UC3 TIPS

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.

Access for partners: YES
External access: NO. External access is not allowed for AP-HP, because they must inform patients about all data processes done with their data.
APHP Paul Brousse
Storage type: Federated
Use cases: UC3 TIPS, UC3 LT
Access for partners: YES
External access: NO. External access is not allowed for AP-HP, because they must inform patients about all data processes done with their data.
APHP Henri Mondor
Storage type: Federated
Use cases: UC4
Access for partners: YES
External access: NO. External access is not allowed for AP-HP, because they must inform patients about all data processes done with their data.
CUSL
Storage type: Central
Use cases: UC1, UC3 TIPS, UC4
Access for partners: YES
External access: Restricted
MUV
Storage type: Central
Use cases: UC1, UC3 TIPS
Access for partners: YES
External access: Restricted
IMPERIAL
Storage type: Central
Use cases: UC1, Controls
Access for partners: YES
External access: Restricted
ULS
Storage type: Central
Use cases: UC1
Access for partners: YES
External access: Restricted

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.

7.2 Models generated in ARTEMIS

7.2.1 Summary of the ARTEMIS VHT models

The ARTEMIS VHT models will combine mechanistic and AI-based modelling to simulate interactions between the liver and the heart across the MASLD disease continuum. All modelling activities require of multimodal patient data, although the volume of data needed varies depending on the approach with mechanistic models typically requiring less patients with an excellent characterisation of the parameters involved in the process under simulation. In addition to the retrospective cases being gathered, some small prospective studies will be conducted to gather rich patient data needed for UC2 and UC3 models.

The VHT models developed in ARTEMIS are intellectual outputs owned according to the rules defined in the project GA and CA, unlike datasets, whose ownership ultimately belongs to the patients, and for which the legal provisions of the project are just governing strict safeguards for their storage (managed by Data Holders) and legitimate use (managed by Data Users) The VHT model developers will own their generated results. They will make their preferred decisions for promoting the long-term usability of their models and fulfil their obligation to use their best efforts during the project and for four years following the end of the grant, to exploit their generated results.

The VHT models span multiple scales, from organ-level haemodynamics to cellular and tissue-level fibrosis processes, and include standalone modules and integrative components. Consistent versioning, appropriate documentation and preferred use of Open-Source licenses will facilitate the reuse of these models by the VHT community. The active role of INRIA in the VHT initiative ensure access to updated information on the status of the VHT Advanced Platform developments and thus alignment with its technical requirements for future interoperability.

7.2.2 Mapping of models and partners positions

An initial mapping of the main models under development is presented in Table 20, including their general purpose, related use cases and preliminary position regarding access. In D9.4 this map will be updated with the final results reported by WP5 and WP6. The license finally selected for each of them will be documented.

Table 20. Map of VHT modellers

INRIA. Vignon-Clementel team
Model: Hepatic haemodynamics model
Related Use Case: UC3
Preliminary access position: Open-source preferred, (e.g., CC-BY or GPL), possibly with protected components if needed for exploitation.
INRIA. Drasdo team
Models: Tissue-level mechanistic model of fibrosis progression
Related Use Case: UC1, UC2
Preliminary access position: Open-source preferred, (e.g., CC-BY or GPL), possibly extending the COMPUTIX library of modelling tools (https://bio.tools/computix).

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.

DKFZ. Klingmüller group
Models: Systems biology models of liver metabolism and inflammation
Related Use Case: UC1, UC2
Preliminary access position: Open-source preferred, (e.g., CC-BY, MIT), possibly with protected components if needed for exploitation.
ULEI. Höhme group
Models: AI-based disease progression and risk assessment models, agent-based models of liver regeneration and fibrosis.
Related Use Case: UC1, UC2, UC4
Preliminary access position: Open-source preferred, (e.g., CC-BY), possibly with protected components if needed for exploitation.
ALU-FR. Timmer group
Models: Statistical/AI models for MASLD progression prediction
Related Use Case: UC1, UC4
Preliminary access position: Open-source preferred, possibly with standard open licences (CC-BY, MIT).
BU. Bouchachia group
Models: AI models for MASLD patient clustering
Related Use Case: UC1, UC2
Preliminary access position: Open-source preferred, possibly with standard open licences (CC-BY).

7.3 Sustainability strategy for ARTEMIS model

7.3.1 Principles

The sustainability of ARTEMIS models is guided by the principles of FAIR, security and transparency. All models should follow clear documentation standards (model description, assumptions, parameters), maintain versioning and include metadata enabling future integration, reuse or extension.

7.3.2 VHT advanced platform

The ARTEMIS Description of Action commits the consortium to ensuring long-term availability of results by aligning with the European Commission's strategy for Virtual Human Twins. Initially, this was expected to occur through the EDITH platform.

Since then, the EDITH initiative has been discontinued and an alternative platform is being promoted by the EC: the **VHT Platform**, designed as a **distributed platform** enabling users to develop, test and integrate VHT models.

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.

In 2025, DG CNECT launched a public procurement for the **Advanced Platform for VHT models** with a budget of €24M.[103] As to date, three consortia comprising organisations such as Ernst & Young, NTT Data and Boston Consulting, have been announced as winners. A separate procurement for a Project Support Office with a budget of €0.5M is ongoing and will support DG CNECT in supervising delivery of the platform.

Although official specifications remain under development, early information highlights several core features for the Advanced Platform for VHT:

- Collaborative, community-driven workspace: for integration and co-simulation of VHT models.
- Distributed and flexible architecture: relying on open-source specifications.
- Access to cutting-edge digital capabilities: including HPC, cloud, edge computing and AI.
- Secure, privacy-preserving environment: with IPR protection and anonymous or synthetic data.
- Federated repository of VHT resources: open-source toolkits and simulation tools.

ARTEMIS´s modellers will actively monitor the development of the VHT Advanced Platform to ensure that their models remain interoperable with the emerging European VHT ecosystem. Model owners are aware of the importance of maintaining their models´ metadata, documentation and interfaces consistent with forthcoming platform specifications, in order to consolidate as contributing players of the European VHT initiative.

7.3.3 Transition to VHT advanced platform

The procurement documentation indicates a 48-month development timeline for the Advanced VHT Platform, with an undefined start date. This implies that ARTEMIS must plan an **intermediate transition period** between project completion and the availability of the new platform. During this period, models will need temporary hosting on stable infrastructures (e.g., institutional Git repositories, Zenodo, or consortium GitLab).

To ensure that models designated for open access are fully FAIR, developers should:

- Provide full documentation (purpose, inputs, outputs, assumptions, mathematical formulations).
- Use standard packaging (e.g. Docker containers, Python packages).
- Ensure testability through validation datasets and example scripts.

  Maintain version control.
- Adopt community standards, e.g.:

  Drasdo’s team often uses repositories like GitHub and model exchange standards (SBML).

  Höhme’s and Timmer’s groups frequently publish computational models on Zenodo and institutional Git servers.

  Vignon-Clementel's group regularly releases code via open academic repositories or INRIA’s GitLab.

  Klingmüller’s group often uses SBML and FAIRDOMHub or similar infrastructures.

[103] Call for tenders “Platform for Advanced Virtual Human Twin (VHT) Models” EC-CNECT/LUX/2024/OP/0014

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.

Each model owner will be responsible for establishing contact with the VHT Advanced Platform team once technical specifications become available.

7.4 Sustainability strategy for ARTEMIS datasets

7.4.1 Principles

The sustainability of the ARTEMIS cohort follows the principles of FAIR data management, while respecting the constraints of the applicable legal frameworks including the DSA. Metadata must be openly available, access procedures must be transparent and all data access must occur within secure, auditable environments. The sustainability strategy prioritises the use of European repositories that ensure Data Holders retain control over their contributions, and allows for reuse by the research community.

7.4.2 Sustainability actions

7.4.2.1 Ensure FAIR metadata and consider the use of DOIs:

Ensuring FAIRness of metadata is under the responsibility of the data curator (MEDEX-BCP). Each dataset will have a title and metadata descriptors. The metadata descriptors of key aspects that make the dataset findable by its right potential users, such as (illustrative, not exhaustive description)

- Version number
- Creation date
- Description
- Subjects count
- Age range
- Sex (# subject)
- Diagnosis (# subject)
- Type of data (clinical, imaging, digital pathology)
- Year of diagnosis range

The conditions of Use or License under which the dataset is offered will also be indicated. The use of Creative Commons Attribution license[104] is recommended as it allows re-distribution and re-use of a licensed work on the condition that the creator is appropriately credited.

The possibility to generate a Digital Object Identifier (DOI) for the ARTEMIS dataset will be assessed since these unique identifiers are very interesting for facilitating the citation of the datasets in scientific publications. A possibility is to use Zenodo to register a DOI for each dataset (free of charge) by providing the necessary metadata. Once published, the metadata of the Zenodo datasets is always publicly accessible, and it can be configured to make the files Public or Restricted to specified users.

[104] https://creativecommons.org/licenses/by/4.0/legalcode

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.

7.4.3 Long-term storage and access management

The ARTEMIS datasets are valuable assets corresponding to anonymised, multimodal patient data that has been harmonised, curated and validated as part of the ARTEMIS project activities for the development of VHT models for the liver-heart axis in a specific use case. In line with the Open Science principles of Horizon Europe and the European Health Data Space vision of facilitating the use of patient data for research and innovation, as stated in the GA, the consortium is committed to facilitate the preservation of the ARTEMIS cohort after the project ends. Ensuring long-term access is essential not only to safeguard the scientific and financial investment made in establishing the ARTEMIS cohort, also to maximise its scientific value. The long-term access will enable future validation studies and replication of findings, as well as development of new models. For the clinical partners, it will also strengthen their institutional visibility and have the reputational benefits of contributing to European research and scientific advancement.

7.4.3.1 Federated datasets

Federated datasets remain physically stored at each Data Holder site. This implies that long-term sustainability of data access typically implies storage costs for the ARTEMIs curated/harmoinised dataset being accessible through the ARTEMIS platform, normally in duplication to the original raw data. Furthermore, federated access implies in-site computational workload for analytics. Each site will be invited to estimate their annual costs and report them in D9.4, together with a realistic commitment for maintaining their ARTEMIS dataset accessible for a defined period after project completion.

As a federated site, each be responsible for keeping access for the cloud to their node open. The access-request process (preliminary described in 7.1.4.1) is planned to be common for central and federated datasets, the designated representative of the involved Data Holders would be voluntarily involved in the access request evaluation.

It is anticipated that, if by the end of the project the sites can demonstrate that they are taking steps toward making their data available for secondary use through the European Health Data Space framework, no additional mechanism would likely be required specifically for the ARTEMIS datasets. However, this will depend on the pace of EHDS implementation at each institution. The intention is to avoid duplication of costs as much as possible and not to require ARTEMIS specific measures if a general framework can already achieve the intended purpose.

In exceptional cases where a federated site is not willing to maintain long-term access to its dataset, it will be asked to provide a clear justification which will be documented in D9.4.

7.4.3.2 Centralised dataset

The annual cost associated with the storage and management of the central cloud datasets will be estimated by BCP in the next period to the extent at that point possible (future storage size and cost development of cloud infrastructure provider will affect actual costs), together with the cost associated to the maintenance of the public metadata catalogue (both central and federated) for the ARTEMIS datasets.

In D9.4 and based on this assessment, BCP will determine a realistic commitment for maintaining the ARTEMIS cloud environment operational for a defined period after project completion. In the next period, the consortium will agree to a minimum continuity period at least for data access for ARTEMIS partners if needed for completion of analyses and publications after the finalisation of the project. Beyond this timeframe, long-term sustainability of access to the ARTEMIS dataset collection will need to be ensured through other options.

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.

7.4.3.3 Options for long-term sustainability of data access

The long-term sustainability of data access may be achieved through:

- Scenario 1: The voluntary commitment of a member of ARTEMIS consortium to assume the long-term management of the ARTEMIS Cohort and the hosting of the centralised datasets.
- Scenario 2: The integration into an established European data-sharing infrastructure that adheres to the highest security and compliance standards.

For scenario 1, if a member of ARTEMIS Consortium expresses an interest in assuming the operational and financial responsibility for hosting and maintaining of the ARTEMIS Cohort, WP9 will document the organisation's strategic priorities justifying such an interest, their proposed technical and operational approach to ensuring secure management, and their previous experience in managing patient datasets. This information will be the subject of an internal assessment by all data holders.

For scenario 2, the possibility of using the VHT Advanced Platform for clinical datasets will be evaluated once the tender winners have begun deploying the platform. However, we understand that the VHT Advanced Platform is intended for sharing modelling data only. In that case, other alternatives will be considered for sharing ARTEMIS datasets under a restricted access modality and with the option for data holders to participate in evaluating access requests. Table 21 presents a preliminary analysis of some options.

Table 21. Options for long-term sustainability of ARTEMIS datasets

Repository / Initiative	Description
ZENODO https://zenodo.org/	The European reference repository for open access
	PROS: Free of charge. Useful for creating DOIs for assets
	CONS: Not specific to health data. Open access, is not designed for controlled access with access request management.
Cancer Image Europe pilot infrastructure (**EUCAIM**, https://cancerimage.eu/),	Pilot repository in production phase, for the Cancer Image Europe initiative
	PROS: Several ARTEMIS partners are founding partners of this initiative. It allows for imaging data and related clinical and lab data. hybrid (federated/central) platform with access request management and secure processing environment.
	CONS: Currently it is specific for cancer, only suitable for UC4 datasets.
The European foundation for the study of Chronic Liver Failure (**EF-CLIF**, https://efclif.com/)	EF-CLIF is a private foundation but with a public duty to conduct its affairs in a transparent manner. They fund large scale observational studies, with +120 participant centres across Europe. The foundation is a participant in on-going Horizon projects such as LEOPARD and DECISION.
	PROS: EF-CLIF manages a restricted access Data Hub which offers access for rearchers to pseudoanonymized clinical and omics data, and biological samples from their large observational studies. Researchers can submit an access request which is assessed by the Sample and Data Usage Committee (SDUC).

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.

	CONS: The Data Hub hosts datasets gathered in studies sponsored by EF-CLIF. If the ARTEMIS consortium is interested in exploring this possibility, conversations should be started to ascertain their level of interest and the conditions they would require.
CANDLE https://candle-project.eu/ or UNCAN.eu (HE projects) https://cordis.europa.eu/project/id/101215206	These are the 2 projects funded under the call HORIZON-MISS-2024-CANCER-01-01 - Use cases for the UNCAN.eu research data platform. CANDLE is aimed at creating national nodes for cross-border data sharing, while UNCAN will establish a federated data hub for cancer research. The projects will finalise in 2030 but the data sharing platforms are aimed to last and facilitate health data sharing.
	PROS: Very ambitious initiatives, called to create the future European data sharing system for multimodal cancer data.
Physionet (https://physionet.org) **Vivli** (https://vivli.org)	PROS: Data sharing platforms specific to biomedical data. Both operate under controlled access.
	CONS: Initiatives hosted in the USA. ARTEMIS consortium will prioritise European initiatives for data sharing.

In the final sustainability plan, the Consortium will examine the most relevant of these options and other restricted-access alternatives in more detail, as well as monitoring any new ones that may arise in the near future. Using any of these initiatives for sharing ARTEMIS datasets will require acceptance of their access governance and may require explicit approval by each data holder´s institutional ethics committee.

8 CONCLUSION

This deliverable contains a preliminary roadmap for the exploitation of ARTEMIS results and for the sustainability of the project's data and model assets. It confirms that the expected KERs are progressing well and that there is clear innovation and market potential. At this stage, the propositions described remain exploratory and depend on the final scope of the delivered results. Therefore, the detailed exploitation roadmap for the most promising results and the sustainability mechanisms for the FAIR datasets and models, will be reassessed with the consortium during the second half of the project.

The next iterations will lead to a final selection of the exploitable assets for which well-defined exploitation routes and planned or in place IPR protection measures will be described. They will also lead to an agreement on the sustainability measures for the FAIR datasets and models- The final strategy will promote that the most valuable ARTEMIS outcomes continue to generate impact beyond the project's lifetime.

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.

9 FUTURE WORK

9.1 Next steps in the KER exploitation strategy

- Expand detailed assessments to other KERs as their development matures and their related deliverables are informed.
- Continue monitoring the undertaken IPR protection actions and support it upon request of the partners' legal / technology-transfer departments, particularly in the case of co-ownership.
- Concentrate efforts on those KERs with the strongest market potential and clear owner commitment, which will be consider top-priority.
- Develop preliminary business strategies for the top-priority KERs, including routes to market or adoption in research, with identification of potential adopters and key alliances to reach market readiness.
- Inform the IPR protection means for the KERs and the business strategies for the top priority KERs in D9.4.

9.2 Next steps in the sustainability strategy

- Define the Access request and the Request evaluation procedures for the datasets and incorporate them as Annex to D9.4: these procedures will be adopted by the centralised repository and recommended for federated sites.
- Confirm the partners positions on their datasets or models sharing: each partner must validate their long-term access conditions to inform D9.4.
- Data Holders initiate any needed local Ethics Committee amendments for legal clearance of anonymised datasets reuse under restricted access conditions.
- Ensure data FAIRness for the datasets eligible for reuse: Complete metadata, register DOIs through metadata upload to Zenodo (optional), register in the public metadata catalogue to be accessible through the ARTEMIs website.
- Integrate all sustainability decisions into the final DMP (D2.4): related to governance, FAIR, legal and technical sustainability elements.
- Prepare model transition plans toward the VHT Advanced Platform: gather latest information on the platform deployment and align the models' documentation and technical interfaces.

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.