



**DARE**  
DIGITAL LIFELONG PREVENTION

CODE NO. PNC0000002

**Spoke 3 Deliverable**

**SP3.D5.2**

# Models preliminary assessment and validation

This research is co-funded by the Ministry of University and Research within the Complementary National Plan PNC-I.1 "Research initiatives for innovative technologies and pathways in the health and welfare sector"  
D.D. 931 of 06/06/2022, PNC0000002 DARE - Digital Lifelong Prevention



## SP3.D5.2. Models preliminary assessment and validation

Deliverable information	
Spoke number and title	Spoke 3 - Digitally-Enabled Secondary And Tertiary Prevention
WP number and title	WP 5 - Digital Tools for Screening and Early Diagnosis
Related task(s)	Task 5.1, Task 5.2, Task 5.3, Task 5.4, Task 5.5, Task 5.6
Lead beneficiary	UNIPD
Contributing beneficiaries	UNIBA, UNIROMA2, IRCSS-AOUBO, ASL BA, PTV
Dissemination level	Public, fully open
Due date	14/6/2025
Actual date of delivery	21/7/2025
Author(s)	Andrea Facchinetti (UNIPD), Giacomo Cappon (UNIPD), Giovanni Sparacino (UNIPD)
Contributors	Fernando Rizzello, Marco Salice, Isadora Beghetti, Maria Grazia Capretti (IRCSS-AOUBO), Loreto Gesualdo, Paola Pontrelli (UNIBA), Giuseppe Liotta, Daniele Di Giovanni, Stefano Cianfarani (UNIROMA2), Giacomo Cappon, Andrea Facchinetti, Giovanni Sparacino, Luca Vedovelli, Dario Gregori (UNIPD)
Quality Assurance	Mattia Veronese (UNIPD), Alessandro Chiarini (Bi-Rex)

## Document history

Version	Date	Author(s) /Reviewer(s) (Beneficiary)	Description
0.1	31/05/2025	Andrea Facchinetti (UNIPD) Giacomo Cappon (UNIPD)	First draft
0.2	04/06/2025	All Pilot Leaders	Additions to all sections
0.3	06/06/2025	Andrea Facchinetti (UNIPD) Giacomo Cappon (UNIPD)	Consolidating contents
0.4	07/06/2025	Giovanni Sparacino (UNIPD)	Revision
0.5	09/06/2025	Giacomo Cappon (UNIPD)	Revision
1.0	10/06/2025	Andrea Facchinetti (UNIPD)	Final document for external revision
1.1	03/07/2025	Mattia Veronese (UNIPD)	Revision
1.2	07/07/2025	Alessandro Chiarini (Bi-Rex)	Revision
1.3	10/07/2025	All Pilot Leaders	Answers to revisions and additions to all sections
2.0	21/07/2025	Andrea Facchinetti (UNIPD) Giacomo Cappon (UNIPD)	Consolidating contents and creations of final document

## Disclaimer

This publication reflects only the author's views and the Funding Agency is not liable for any use that may be made of the information contained therein.

Table of contents

<b>1. Publishable summary.....</b>	<b>7</b>
<b>2. Introduction.....</b>	<b>8</b>
<b>2.1. WP5 overview and aims .....</b>	<b>8</b>
<b>2.2. Overview of the project status at month 30.....</b>	<b>8</b>
<b>2.3. Deliverable outline .....</b>	<b>9</b>
<b>3. Progress and Preliminary Validation of Pilot Studies .....</b>	<b>10</b>
<b>3.1. Task 5.1 – IBD Care Through Hub&amp;Spoke Infrastructure (responsible: IRCSS-AOUBO).....</b>	<b>10</b>
3.1.1. Overview.....	10
3.1.2. Data used for preliminary assessment and validation .....	10
3.1.3. Devices employed in the study .....	11
3.1.4. Developed models/algorithm/platforms .....	12
3.1.5. Evaluation metrics and statistical analysis .....	14
3.1.6. Results and discussion .....	15
3.1.7. Next steps .....	16
<b>3.2. Task 5.2 – Home monitoring-based approach to support diagnostic, therapeutic and assistance pathway of patients with chronic kidney disease (responsible: UNIBA) .....</b>	<b>17</b>
3.2.1. Overview.....	17
3.2.2. Data used for preliminary assessment and validation .....	17
3.2.3. Devices employed in the study .....	18
3.2.4. Developed models/algorithm/platforms .....	19
3.2.5. Evaluation metrics and statistical analysis .....	19
3.2.6. Results and discussion .....	20
3.2.7. Next steps .....	20

<b>3.3. Task 5.3 – Prevention of adverse events in preterm and term infants by remote monitoring (responsible: IRCSS-AOUBO)</b>	<b>22</b>
3.3.1. Overview	22
3.3.2. Data used for preliminary assessment and validation	22
3.3.3. Devices employed in the study	23
3.3.4. Developed models/algorithm/platforms	24
3.3.5. Evaluation metrics and statistical analysis	25
3.3.6. Results and discussion	26
3.3.7. Next steps	26
<b>3.4. Task 5.4 – Therapy optimization and prevention of adverse events in diabetes management (responsible: UNIROMA2, co-responsible: UNIPD)</b>	<b>28</b>
3.4.1. Overview	28
3.4.2. Data used for preliminary assessment and validation	28
3.4.3. Devices employed in the study	29
3.4.4. Developed models/algorithm/platforms	30
3.4.5. Evaluation metrics and statistical analysis	33
3.4.6. Results and discussion	34
3.4.7. Next steps	35
<b>3.5. Task 5.5 – Non-medical wearable devices for monitoring caloric intake (responsible: UNIPD)</b>	<b>37</b>
3.5.1. Overview	37
3.5.2. Data used for preliminary assessment and validation	37
3.5.3. Devices employed in the study	38
3.5.4. Developed models/algorithm/platforms	39
3.5.5. Evaluation metrics and statistical analysis	41
3.5.6. Results and discussion	42
3.5.7. Next steps	42

<b>3.6. Task 5.6 – Prevention/mitigation of frailty in the continuum of care framework (responsible: UNIROMA2)</b> .....	<b>43</b>
3.6.1. Overview.....	43
3.6.1. Data used for preliminary assessment and validation .....	44
3.6.2. Devices employed in the study .....	44
3.6.3. Developed models/algorithm/platforms .....	44
3.6.4. Evaluation metrics and statistical analysis .....	46
3.6.5. Results and discussion.....	47
3.6.6. Next steps.....	48
<b>4. Conclusion and next phase</b> .....	<b>49</b>

## 1. Publishable summary

This section offers a succinct outline of the content of deliverable S3.D5.2, including objectives, methodologies employed, key findings and their implications. While, for confidentiality reasons the deliverable is flagged with 'restricted access,' the present summary is designed to be publishable, ensuring that relevant insights and highlights on the results provided within the DARE initiative can be timely shared with a wide audience.

In Spoke 3 of the DARE project, WP 5, titled “Continuity of care interventions for secondary and tertiary prevention,” focuses on developing new Information and Communication Technologies (ICT) solutions for continuous care interventions. WP5 comprises six tasks which integrate various technological tools, from wearable devices to algorithms supporting clinical decisions, with different clinical research methodologies, including observational and randomized study designs. Six pilot studies (one for each task) are ongoing. They involve diverse patient groups, such as infants, older adults, diabetics, Inflammatory Bowel Disease (IBD) patients, and chronic kidney disease patients. The solutions developed in the DARE project build upon the experiences of the responsible research groups and aim to open new clinical and prevention pathways, ultimately improving patient quality of life and the effectiveness of health services.

A first deliverable, handling concept and relevant design of the models, was produced for WP5 at month 12. During the second phase of the activities, between month 12 and month 30, significant progress has been made in the development and initial assessment of the digital tools for screening and early diagnosis. Prototypes have been deployed in real-world settings, and patient enrollment has either begun or is imminent. Preliminary data collection and stress testing have informed refinements in system interfaces, clinical reporting, alerting logic, and backend integration. The six pilot studies are designed not only to improve care quality and decision-making but also to enhance patient engagement and reduce unnecessary hospital use.

The present deliverable report, for each task of the WP, the current state of implementation, the initial results, and the planned next steps.

## 2. Introduction

### 2.1.WP5 overview and aims

WP 5 aims to design, develop, and validate ICT-enabled interventions that support continuity of care and early diagnosis, as part of a broader effort to enable secondary and tertiary prevention. It encompasses six tasks, listed in Table 1 with the units involved and the PIs.

Task	Responsible	Pilot study	PI
5.1 IBD care through hub&spoke infrastructure	IRCCS AOU BO	IBD care through hub&spoke infrastructure	Marco Salice
5.2 Home monitoring – based approach to support diagnostic, therapeutic and assistance pathway of patients with chronic kidney disease	UNIBA	Digital therapy and telemedicine approaches for nutritional intervention in ADPKD	Loreto Gesualdo
5.3 Prevention of adverse events in preterm and term infants by remote monitoring	IRCCS AOU BO	NICU at home: A pilot study to promote family-centred care via telemedicine and e-health to prevent health issues in preterm and term infants	Maria Grazia Capretti
5.4 Therapy optimization and prevention of adverse events in diabetes management	UNIROMA2, UNIPD	A digital, integrated and scalable mobile platform based on wearable sensors to prevent/reduce the risk of complications in type 1 diabetes	Stefano Cianfarani, Andrea Facchinetti
5.5 Non-medical wearable devices for monitoring caloric intake	UNIPD	FITMATE: Food Intake Tracker and Metabolic Evaluation for Health Enhancement	Dario Gregori
5.6 Prevention/mitigation of frailty in the continuum of care framework	UNIROMA2	Developing of a social and health care integrated model to reduce the overcrowding of Emergency Room and inappropriate hospital admissions	Giuseppe Liotta

Figure 1.1.1: List of the six tasks of WP5 and titles of the ongoing pilot studies.

Each task embodies a dedicated pilot study. The pilot studies target various patient populations such as preterm infants, patients with chronic conditions (e.g., Inflammatory Bowel Disease - IBD, Autosomal Dominant Polycystic Kidney Disease - ADPKD, diabetes), and the frail elderly. Each pilot study integrates digital solutions, ranging from wearable sensors and Internet of Things (IoT) devices to Artificial Intelligence (AI)-driven platforms and mobile applications, into patient care protocols. These tools are designed to support clinical decision-making, enable remote monitoring, and improve the quality and personalization of care. The six pilot studies also serve as testbeds for evaluating the flexibility, acceptability, and efficacy of ICT tools in diverse healthcare environments.

## 2.2. Overview of the project status at month 30

As of month 30, all six pilot studies have progressed significantly from their initial design phase:

- Digital platforms and devices have been deployed or are in advanced development stage.
- Early clinical testing, data integration, and usability assessments of tools have been or is close to be completed.
- Patient recruitment has begun or is about to begin in line with approved study protocols.
- Backend analytics modules and decision support algorithms are undergoing iterative testing and refinement based on pilot data.
- The pilots are entering a critical phase of full-scale validation, where longitudinal data collection will support the final statistical evaluation of model performance and health outcomes.

## 2.3. Deliverable outline

After the overview of WP5 general goals given in the present section, the following section (Section 2) provides a comprehensive report of the progress made in each of the six tasks of WP5, describing specific aims, data used, tools developed, validation strategies, intermediate results, and future steps. Section 3 presents overarching conclusions and highlights key milestones and challenges anticipated for the upcoming project phase (months 30–48).

### 3. Progress and Preliminary Validation of Pilot Studies

#### 3.1.Task 5.1 – IBD Care Through Hub&Spoke Infrastructure (responsible: IRCSS-AOUBO)

##### 3.1.1. Overview

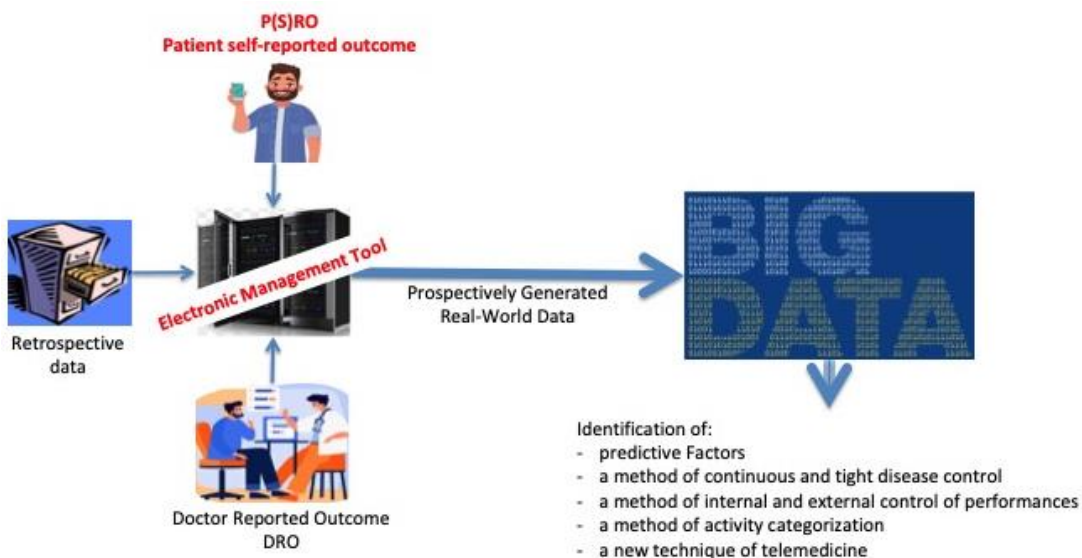


Figure 2.1.1: Graphical abstract of Task 5.1.

The Emilia-Romagna Inflammatory Bowel Disease (IBD) hub-and-spoke infrastructure is being expanded to enhance quality of care monitoring, facilitate clinical research, and enable real-time clinical decision support. Task 5.1 focuses on the development and validation of a digital tool designed to support structured data entry during clinical visits and enable patient self-reporting through an integrated mobile application. This system will be embedded within the regional IBD network to collect real-world data, produce automated clinical summaries, and support performance benchmarking.

By integrating retrospective and prospective data streams, the digital tool will enable predictive analytics and timely identification of patients at risk of disease worsening or treatment failure. Ultimately, the project aims to support early intervention, reduce variability across centers, and improve patient outcomes.

Initial real-world deployment has begun, and further enhancements are in progress based on clinician feedback and system performance evaluations.

### 3.1.2. Data used for preliminary assessment and validation

The digital tool is under evaluation through a pilot implementation at the IRCCS University Hospital of Bologna – Policlinico di Sant'Orsola (IRCCS-AOUBO). Integration into routine outpatient care facilitates both retrospective and prospective data collection. Historical data are obtained from existing patient records, while new data are gathered in real time during clinical encounters via the digital platform.

The dataset comprises structured clinical information, laboratory results, endoscopic findings, imaging reports, and pathology data for a cohort potentially encompassing up to 10,000 patients with IBD, reflecting the patient population managed at the center and within the Emilia-Romagna region. Informed consent has been provided by over 1,000 patients to date. Data entry is standardized utilizing validated disease activity indices, including the Mayo clinical and endoscopic score for ulcerative colitis, Montreal classification, Harvey Bradshaw Index for Crohn's Disease, Simple Endoscopic Score for Crohn's Disease (SES-CD), Limberg score, and IBD-Q.

Real-world data collection commenced following the approval of a protocol amendment by the local ethics committee in late 2024 and the deployment of the beta version of the digital tool in mid-2025. The tool enables the analysis of disease trajectories, treatment responses, and center-level performance metrics. Additionally, a mobile application currently in development aims to allow patients to input periodic laboratory values, symptom data, and patient-reported outcomes, thereby enhancing data continuity and patient engagement.

### 3.1.3. Devices employed in the study

<b>DEVICE#1</b>	
<b>Model</b>	Custom digital tool for IBD management, developed by IRCCS-AOUBO in collaboration with software partners.
<b>What is it measuring? What is the role of the device in the study?</b>	The digital tool supports structured collection of clinical, laboratory, endoscopic, and imaging data during outpatient visits. It enables clinicians to enter patient data in real time and automatically generates standardized medical reports. Data entry is event-based, occurring during clinical interactions.

	<p>The tool serves as the operational backbone of the project, integrating data collection, clinical reporting, and backend analytics. It allows clinicians to input standardized information during visits while simultaneously producing patient-facing summaries. The system is being progressively integrated with hospital infrastructure for laboratory, endoscopic, and imaging data, and includes modules for alert generation and future patient self-reporting. The digital tool will also incorporate a patient-facing application currently under development. This app will allow patients to input periodic laboratory results, symptom diaries, and validated patient-reported outcomes (PROs), further enhancing the continuity and granularity of real-time data capture.</p>
<p><b>Total number of devices used in the study</b></p>	<p>Single-instance deployment at the IBD Unit, IRCCS-AOUBO</p>

### 3.1.4. Developed models/algorithm/platforms

The digital tool under development is a modular software system designed to support high-quality IBD care by enabling structured data collection, longitudinal monitoring, automated alerts, standardized triage, and center performance auditing. It includes a desktop interface for clinicians, a patient-facing self-reporting component, and a backend analytics engine.

The clinician-facing component (allows structured entry of clinical, laboratory, imaging, and endoscopic data during patient visits. It aligns with established disease activity indices and enables semi-automatic generation of patient reports. Recent stress-testing of the clinician-facing component revealed system performance issues that are currently being addressed. A second development phase will expand the system's capabilities to include automatic integration with hospital repositories and refinement of clinical documentation features. The patient-facing component (web/phone app) is under early-stage development and will allow patients to self-report symptoms, lab results, and validated PROs via web or mobile interface. This aspect is expected to enhance early detection of clinical deterioration and improve patient engagement.

The screenshot displays a web-based interface for data entry. On the left is a navigation menu with options like 'Anamnesi generale', 'Anamnesi IBD', 'Storico attuale', 'Esami di laboratorio', 'Altra terapia in corso', 'Conclusioni', and 'Terapia consigliata'. The main area is divided into sections: 'Anamnesi generale' with radio buttons for 'Fumo', 'Allergie', 'Intolleranze', 'Patologie Associate', and 'Familiarità di altre patologie'; 'Anamnesi IBD' with radio buttons for 'Familiarità per IBD', 'Esordio dei sintomi', and 'Tipologia di malattia diagnosticata'; and 'Eventi farmacologici' which includes a table of medication events.

ID	Data	Tipologia	Farmaco	Dosaggio	Stato	Azioni
538a07	27/05/2025	Inizio somministrazione farmaco	Metronidazole - FLAGYL 500 cpr (018505038)	250 mg cps, 1 cp x 3 volte al giorno per 7 giorni al mese	3	Nascondi, Cancella, Modifica
82563e	27/02/2025	Inizio somministrazione farmaco	Mesalazine - MESAVANCOL 120 cpr gastrores 1.200 mg rilascio prolungato (037734023)	2 cp al giorno	3	Nascondi, Cancella, Modifica

Figure 2.1.2: Screenshot 1 of the clinician interface of the digital tool showing structured data entry.

The analytics engine is designed to perform trend analysis and to generate alerts when deviations from historical patterns are detected. These will be coded using a green/yellow/red flag system to facilitate triage and intervention decisions.

Co-design sessions with clinicians were used to refine the interface and report formats. All data are stored securely in an encrypted, GDPR-compliant environment with access controlled by user role. The digital tool currently operates independently from the regional Electronic Health Record (HER) system but is designed for future integration. The platform is hosted on secure internal servers.

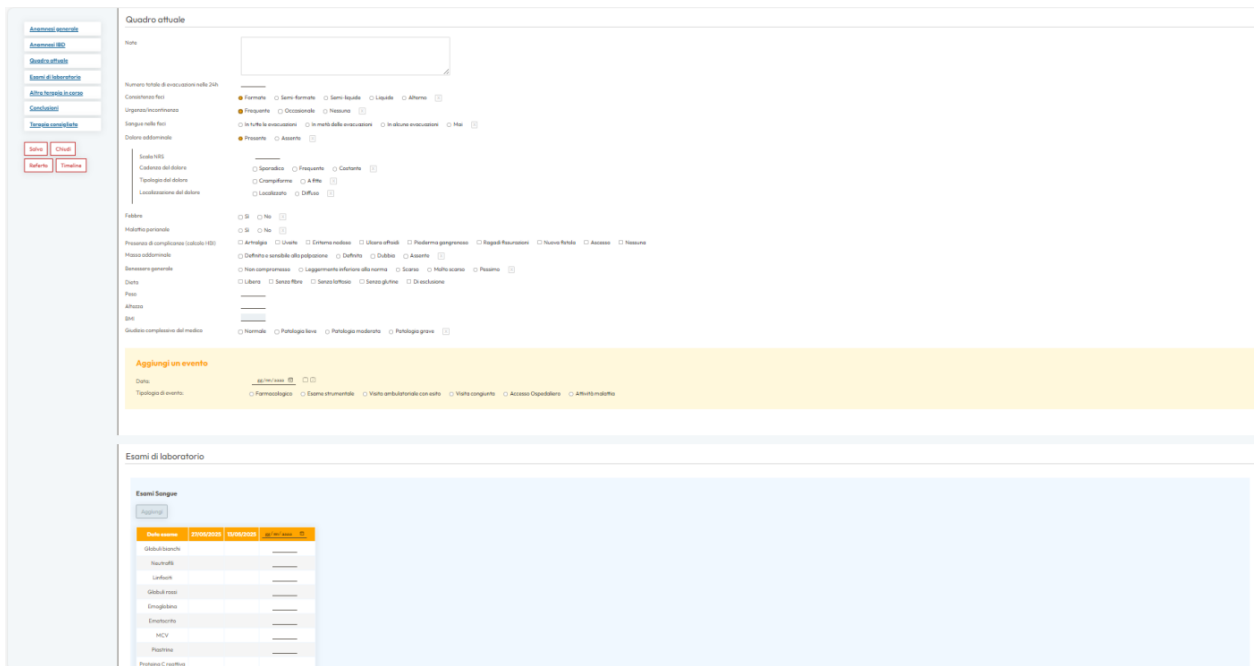


Figure 2.1.3: Screenshot 2 of the clinician interface of the digital tool showing structured data entry.

### 3.1.5. Evaluation metrics and statistical analysis

The validation approach is being carried out in sequential phases. In the initial phase, stress-testing was conducted using simulated clinical scenarios to evaluate system performance, data entry reliability, and the effectiveness of the alerting system. Feedback from these exercises guided revisions to the medical reporting format and alerting thresholds, and revealed software-related issues that are currently being addressed. Usability is being assessed through clinician feedback and task-based walkthroughs. Preliminary evaluation metrics collected during internal testing included data completeness (% of required fields filled) and average time for visit documentation. Future testing will include further metrics such as latency under simulated load and user-reported usability via structured feedback forms.

Statistical validation of predictive functions will begin once adequate volumes of structured data are collected, estimated at approximately 1000 patients with longitudinal follow-up to ensure robust model training and testing. Performance will be assessed using Receiver Operating Characteristic (ROC) curves, sensitivity/specificity calculations, and survival analyses where applicable, with cross-validation techniques employed to avoid overfitting. The predictive models are expected to address clinically relevant questions, such as which patients are at increased risk of relapse following biologic treatment discontinuation, which patients are at increased risk of treatment non-response, can early trends in C-reactive protein, fecal

calprotection, or bowel ultrasound parameters predict disease course, etc. Cross-validation techniques will be used to ensure independence between training and test datasets.

### 3.1.6. Results and discussion

Initial implementation of the digital tool enabled structured clinical documentation and automatic generation of visit summaries, which represent a significant shift from prior unstructured workflows. Through simulated use cases and stress-testing during quarters 7 to 9, the development team identified several areas for technical refinement, including interface responsiveness and alerting performance under high-volume data loads. Many of these issues have been addressed in collaboration with the software developers.

The clinical report format has been revised based on feedback from physicians, enhancing clarity, readability, and adherence to documentation standards. This refinement also improves the utility of the reports for longitudinal patient management and decision-making. Integration efforts have focused on aligning the tool with the hospital's systems for laboratory values, imaging reports, and endoscopy, and are progressing toward full connectivity.

Real-world data entry began in late 2024 following approval of the amended protocol by the Ethics Committee and release of the beta version of the clinician-facing digital tool in mid 2025. Early usage in clinical settings has highlighted the tool's potential to improve data completeness and support timely clinical decision-making. Preliminary results from internal testing demonstrated high data field completion (>90%) in structured entry and <10 minutes average time for generating clinical documentation per patient.

Although the patient self-reporting component has not yet been launched, its specifications have been defined, and development is underway. This module is expected to expand the tool's scope by enabling remote symptom monitoring and patient-reported outcome (PRO) collection, contributing to earlier detection of disease activity and better longitudinal data.

Initial clinician feedback has been generally positive, particularly regarding the streamlined documentation process and potential for real-time decision support. However, some concerns have been raised regarding the time required to complete data entry in complex cases, prompting plans to introduce smart data entry features.

Evaluation of the system's predictive algorithms and triaging logic will be a focus of the next development phase. The alerting logic is being refined with thresholds adjusted to minimize false positives and enhance actionable notifications. Key indicators such as relapse risk, treatment non-response, and early signs of frailty will be modeled using accumulating data. These predictive functions will be evaluated using ROC analysis, time-to-event statistics, and cross-validation once adequate data volume is achieved.

In summary, the digital tool has demonstrated feasibility and early-stage effectiveness in capturing structured clinical data and supporting clinical workflows. Further refinements are ongoing, and real-world integration continues to progress. The pilot supports the vision of a scalable, modular digital platform capable of improving IBD management, enhancing care standardization, and reducing variability across centers and physicians.

### 3.1.7. Next steps

In the coming months, development will focus on completing the mobile app and ensuring full integration with the clinician-facing tool. Patient-facing features will support remote symptom tracking, entry of periodic laboratory values, and completion of standardized PROs.

Integration with laboratory and imaging databases will be finalized, and alerting logic refined based on real-world use. Following this, the platform will be deployed across additional IBD centers in Emilia-Romagna. Predictive modeling functions will be activated, using accumulated longitudinal data to identify early indicators of treatment failure, relapse, and transition to frailty.

Another planned milestone is the activation and validation of a digital triage module designed to support outpatient access prioritization. This triage component will use an adaptive, questionnaire-based algorithm to assign urgency levels based on symptoms, lab results, and treatment status. The tool will be evaluated against clinician judgment and optimized using machine learning techniques. The goal of the triage module is to standardize referral urgency and reduce subjectivity in appointment scheduling, supporting more efficient outpatient care pathways.

Scientific dissemination will include submission of results to conferences and peer-reviewed journals in gastroenterology and digital health. Interoperability standards and regulatory pathways (e.g., CE certification as a medical device) will be explored.



Long-term, the project aims to deliver a modular, scalable digital platform to support chronic disease management in both high-resource and remote care settings, improving outcomes and reducing care variability in IBD.

### 3.2.Task 5.2 – Home monitoring-based approach to support diagnostic, therapeutic and assistance pathway of patients with chronic kidney disease (responsible: UNIBA)

#### 3.2.1. Overview

## RESEARCH NEEDS AND AIMS



## HOW

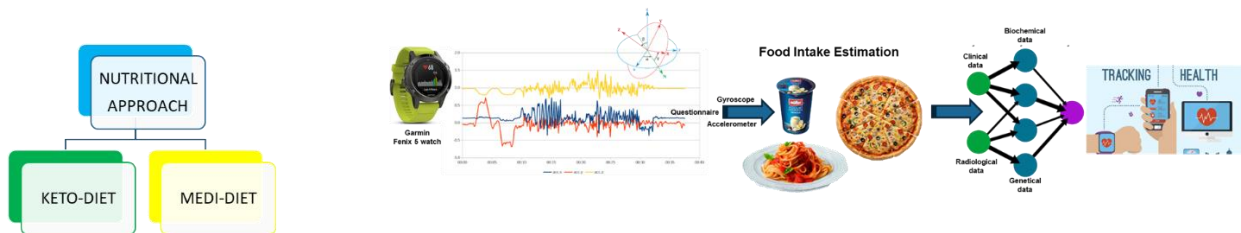


Figure 2.2.1: Graphical abstract of Task 5.2.

The aim is to integrate an IoT infrastructure with clinical, imaging data and genetic background to guide a tailored nutritional (e.g., ketogenic diet) and pharmacological intervention (e.g., metformin vs tolvaptan) for improving patient care in relation to the risk of disease progression in patients with Autosomal Dominant Polycystic Kidney Disease (ADPKD).

IoT technology has the potential to improve the prevention and prediction of end-stage renal disease (ESRD) in patients with ADPKD by providing remote monitoring, early detection of kidney function decline, medication and lifestyle management, also enabling data analytics. Using a wearable device, it will be possible to monitor patients' eating habits, provide an accurate estimate of the quality of diet and caloric/protein intake and analyse the impact of these interventions on kidney function decline and clinical outcomes in patients with ADPKD stratified according to the different clinical and genetic characteristics.

### 3.2.2. Data used for preliminary assessment and validation

Data will be collected by patients. Ethical approval of the study has been obtained on June 13th 2024 and hospital authorization on April 1st 2025. Enrollment will start on June 2025 aiming to enroll 150 patients with Autosomal Dominant Polycystic Kidney Disease [1:400–1:1.000 live births; >12 millions worldwide, about 205.000 people in Europe (EU)], aged 18-60 years with a defined genetic cause of the disease, with baseline eGFR > 30 ml/min/1.73m<sup>2</sup>, randomized to receive a ketogenic diet (high-fat, moderate protein and low-carbohydrate intake based on a ratio of 10:4:1 (in grams) and calorie requirements based on specific patients' condition (at least 20-25 kcal/kg body weight) or a control group with a normal dietary approach. Clinical parameters will be collected every three months for 1 year, Magnetic Resonance Imaging (MRI) volumetry will be done at time 0 and after 6 and 12 months. At time 0 all the patients will receive a specific diet after nutritional counselling and a smartwatch to collect data that will be used to implement a food questionnaire with Redcap interface created in collaboration with UNIPD to collect all the data. Collected data will undergo a standardized preprocessing pipeline including data cleaning, normalization of clinical variables, categorical encoding, and temporal alignment of longitudinal variables.

The final goal is the creation of a Decision Support System for clinicians managing ADPKD patients. To ensure independence between training and validation/test datasets, data from each patient will be kept entirely within one partition. A typical data split will be 80% training, 10% validation, and 10% testing. Stratified cross-validation will also be employed to maintain distributional balance of key variables across folds. This approach avoids temporal and inter-subject data leakage, especially in the longitudinal setting.

At all timepoints stool samples will also be collected to perform microbiome analysis on the two groups with different dietary approaches. Furthermore the 3D Frame platform by Bi-Rex will be tested to develop Learning Technologies approaches for ADPKD patients.

### 3.2.3. Devices employed in the study

<b>DEVICE#1</b>	
<b>Model</b>	FitBit Sense
<b>What is it measuring? What is the role of the device in the study?</b>	The Fitbit Sense is a commercially available smartwatch equipped with multiple sensors, including a 3-axis accelerometer and gyroscope. In this study, it was employed to non-invasively collect motion data from the participant's wrist during daily activities. The recorded features included triaxial

	acceleration, angular velocity (pitch, roll, yaw), signal magnitude vector (SMV), and derived energy metrics. These data streams were used to characterize wrist movement patterns during specific behaviors, particularly food intake, enabling downstream analysis and classification through machine learning approaches.
<b>Total number of devices used in the study</b>	5

### 3.2.4. Developed models/algorithm/platforms

All the models, algorithms and platforms will be developed in collaboration with UNIPD, by using the technologies and the platforms described in Baldi I, Lanera C, Bhuyan MJ, Berchialla P, Vedovelli L, Gregori D. "Classifying Food Items During an Eating Occasion: A Machine Learning Approach with Slope Dynamics for Windowed Kinetic Data. Foods." 2025; 14(2):276.

For this pilot study the raw data related to food intake will be provided by smartwatches in FIT (Flexible and Interoperable Data Transfer) file format. The related FIT-SDK software v.20 development kit provides the FitCSVTool.jar JavaTM program, which can convert FIT files in the common CSV format (Comma Separate Value). Converted raw data will be merged and imported into R, powered by the tidyverse (v.2.0) bundle of packages, and the caret package (v.6.0).

The model used to manage these data will be an implementation of the random forest (RF)-based multiclassification algorithms previously used, and preliminary validated by UNIPD to identify a set of specific food items on meals.

Data related to food intake will also be integrated with clinical and genetic data collected over time from patients to promote the creation of a Decision Support System for clinicians managing ADPKD patients. Genetic data will allow us to divide patients according to the molecular mechanisms responsible of the disease. MRI data will be analysed by radiologist and will give information on the numbers and dimensions of cysts at time 0 and at the follow up in both groups (patients on mediterranean diet and patients on Ketogenic diet). Clinical data will also be collected and analysed. Different Artificial Intelligence (AI) I methods will be used once data will be available for testing and validation. All data will be collected on Redcap and shared among the groups as anonymized samples. The development of the Decision Support System and

associated tools will follow a user-centered design approach, including co-design sessions with clinicians, nutritionists, and patients, to ensure usability and relevance in real clinical settings.

### 3.2.5. Evaluation metrics and statistical analysis

Model Evaluation and Statistical Validation Approaches will be assessed, in collaboration with UNIPD, by using the technologies developed and described within Task 5.5. Once data will be available, the plan is to employ several widely accepted evaluation metrics and validation techniques.

A Cross-Validation (CV) strategy will be used to select the optimal variables and obtain the best model, which avoids the problem of over-fitting. The criterion considered for the selection of the parameter will be the maximization of the global validation accuracy (i.e., the proportion of correctly predicted samples in the validation set). In addition to accuracy, other performance metrics such as precision, recall, F1-score and area under the ROC curve (AUC), will be evaluated depending on the specific model and task (e.g., binary or multiclass classification). These metrics are widely used in clinical decision support and health data analytics literature, and allow for a more nuanced evaluation of classification performance.

To ensure robust evaluation of model performance and minimize overfitting the dataset will be randomly split into training (80%) and test (20%) sets while preserving class distribution (stratified split).

### 3.2.6. Results and discussion

The enrolment of the initial patient cohort is scheduled for June 2025. The objective is to evaluate the effects of a ketogenic dietary approach on disease progression, specifically in terms of total kidney volume increase, cyst number and growth, and the associated decline in renal function over time. Additionally, the study seeks to enhance remote monitoring of ADPKD patients and clinical outcomes through the implementation of specialized wearable devices designed to accurately detect and track patient parameters and habits, including adherence to prescribed dietary and treatment regimens. Specific expected outcomes include:

- Personalized treatment: IoT devices can collect real-time data on an individual's health status, which can be used to personalize treatment plans for ADPKD patients based on their specific needs and disease progression.
- Improved medication adherence: IoT devices can help patients track their medication intake and remind them to take their medication on time, which can improve medication adherence and ultimately lead to better health outcomes.

- Remote monitoring: IoT devices can allow for remote monitoring of ADPKD patients, which can help reduce the number of in-person visits to healthcare facilities, improve patient comfort and satisfaction, and lower healthcare costs.
- Lifestyle modifications: IoT devices can provide patients with real-time feedback on their lifestyle choices, such as diet and exercise, and encourage them to make healthy changes that can improve their overall health and potentially slow the progression of ADPKD.
- Economic outcomes: Remote monitoring with IoT devices may reduce the need for hospitalizations, emergency department visits, and in-person clinic visits, potentially leading to cost savings for patients and healthcare providers.

### 3.2.7. Next steps

The enrollment of patients is scheduled for completion within six months, followed by the collection of follow-up data. All collected data will be utilized to finalize, implement, and test the complete system. Project outcomes are intended to be disseminated at leading international conferences in nephrology, including the European Renal Association (ERA) Congress and the American Society of Nephrology (ASN), as well as in the technology sector at the IEEE Engineering in Medicine and Biology Society (EMBS) - International Conference on Biomedical and Health Informatics (BHI). Furthermore, a final article is planned for submission to internationally indexed journals to describe the clinical trial and highlight the benefits of ketogenic diet approaches for ADPKD patients. An additional manuscript is also planned, detailing the models, algorithms, and platforms developed for the Decision Support System designed for clinicians managing ADPKD patients.

### 3.3.Task 5.3 – Prevention of adverse events in preterm and term infants by remote monitoring (responsible: IRCCS-AOUBO)

#### 3.3.1. Overview

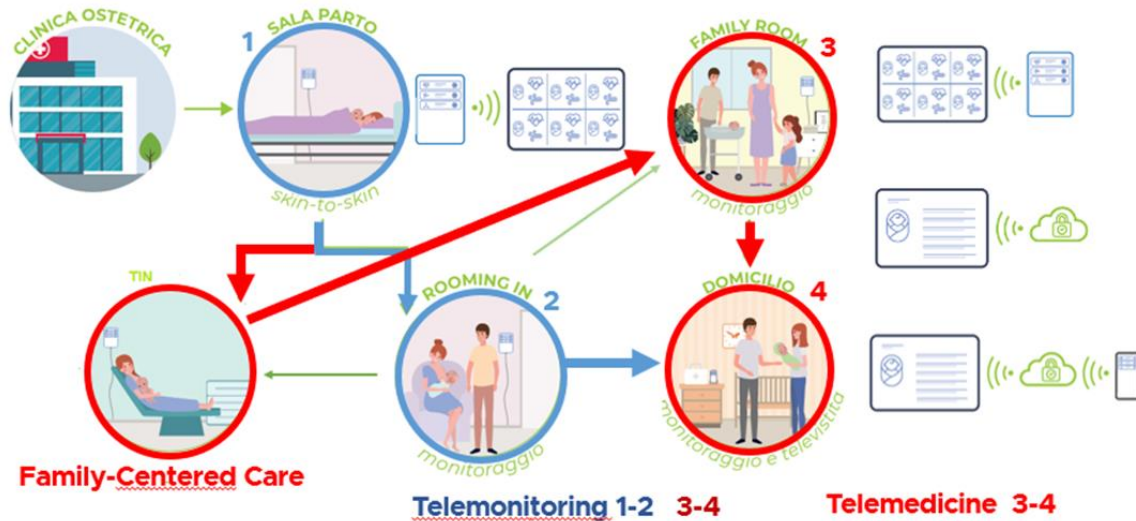


Figure 2.3.1: Graphical abstract of Task 5.3.

This pilot study comprises two distinct clinical trials designed to prevent health issues and improve neonatal care in both preterm and term infants by leveraging telemedicine and e-Health technologies. Pilot A aims to enhance rooming-in safety for term infants and prevent Sudden Unexpected Perinatal Collapse (SUPC). 2,000 infants will be enrolled annually, using wearables to monitor vital signs for 48-72 hours. Data on bradycardia, bradypnea, and apnea will identify SUPC risks and refine rooming-in protocols. Pilot B focuses on enabling the safe and early discharge of preterm infants from the Neonatal Intensive Care Unit (NICU). This trial will involve 40 preterm infants annually. Family-Centered Care principles will be integrated with e-Health technologies, including electronic health records (EHRs), telemonitoring, televisits, and wearable devices, for remote support, empowering families during and after hospitalization. Patients' outcomes will be assessed via parent-reported experience questionnaires and interviews, as well as clinical neonatal outcomes such as length of stay, early re-admission, and emergency care service utilization.

#### 3.3.2. Data used for preliminary assessment and validation

In Pilot A, the study population comprises all infants born at the Delivery Room – Obstetrics and Prenatal Medicine Unit of IRCCS AOU di Bologna with a gestational age of  $\geq 35$  weeks, meeting eligibility criteria for

Skin-to-Skin Contact (SSC) within the first 2 hours of life, and subsequently for rooming-in according to clinical guidelines. The prospective study will enroll approximately 1,000 participants. Enrollment will proceed consecutively following the acquisition of informed consent. Newborns delivered spontaneously and deemed eligible for SSC per clinical guidelines will be equipped with a wearable medical device during the first 2 hours of life in the delivery room. During the subsequent rooming-in period, all newborns, regardless of delivery method (spontaneous or C-section), will utilize the device for the first 48–72 hours of life. Post the initial 2-hour window, monitoring will be conducted intermittently. The monitoring system is designed to non-invasively measure heart rate, respiratory rate, and oxygen saturation. The study aims to quantify SUPC prodromal events (bradycardia, bradypnea, apnea) identified by the ComfTech HOWDY BABY device (ComfTech S.r.l.) during periods not covered by standard clinical monitoring within the newborns' first 72 hours of life. At the end of the monitoring and data acquisition period for each patient, thanks to the electronic device's setup, data collected will be exported to csv/tab format for subsequent analysis.

In Pilot B, the study population includes preterm infants born at <37 weeks of gestation, with low care requirements, who have reached at least 34 weeks corrected age and a minimum weight of 1,800 g, and who are hospitalized at the Neonatology and Neonatal Intensive Care Unit of IRCCS AOU BO. Based on historical data, approximately 40 infants are expected to be enrolled. These participants will be included in a program integrating Family-Centered Care principles with e-Health technologies, encompassing electronic health records, post-discharge real-time audiovisual communication, and remote patient monitoring via wearable devices to facilitate the NICU-to-home transition. Telemedicine consultations and telemonitoring will be conducted with NICU graduates within 15 days post-discharge to identify clinical concerns and deliver continuous caregiver support.

### 3.3.3. Devices employed in the study

DEVICE#1	
<b>Model</b>	Howdy Baby Electronic Unit, Wearable Textile Units
<b>Vendor</b>	Comftech S.r.l
<b>What is it measuring? What is the role of the device in the study?</b>	Heart rate and respiratory rate. Vital sing monitoring.
<b>Total number of devices used in the study</b>	200 wearable reconditionable textile units , 90 reusable Baby Electronic Unit

DEVICE#2	
<b>Model</b>	Pulse oximeter OXY-10
<b>Vendor</b>	GIMA S.p.A.
<b>What is it measuring? What is the role of the device in the study?</b>	Oxygen saturation. Vital sign monitoring.
<b>Total number of devices used in the study</b>	90, reusable

DEVICE#3	
<b>Model</b>	App HOWDY
<b>Vendor</b>	Comftech S.r.l
<b>What is it measuring? What is the role of the device in the study?</b>	It displays the parameters and allows for report creation
<b>Total number of devices used in the study</b>	90

DEVICE#4	
<b>Model</b>	MHP Patient Mobile App and MHP Platform
<b>Vendor</b>	MEDIACLINICS S.r.l
<b>What is it measuring? What is the role of the device in the study?</b>	Direct and remote visualization of recorded physiological parameters, and audio-video calls between discharged patients and healthcare staff
<b>Total number of devices used in the study</b>	4

### 3.3.4. Developed models/algorithm/platforms

The study employs a combination of advanced devices and platforms to monitor and support both term and preterm infants.

For Pilot A, the primary monitoring system is the ComfTech HOWDY BABY Device, designed for newborns. This system includes an electronic unit responsible for data acquisition and transmission, complemented by wearable textile units with integrated sensors in comfortable fabrics for non-invasive, continuous monitoring. Additionally, a Pulse Oximeter OXY-10 (GIMA S.p.A.) is used to measure oxygen saturation (SpO<sub>2</sub>), a critical vital sign. The HOWDY App (ComfTech S.r.l.) functions as the interface for healthcare professionals, providing real-time visualization of monitored parameters and facilitating detailed report generation.

Pilot B, focusing on preterm infants, utilizes the MHP Patient Mobile App for parents and caregivers, alongside the MHP Platform for clinicians and researchers (both provided by MEDIACLINICS S.r.l.). This comprehensive platform supports the transition from the Neonatal Intensive Care Unit (NICU) to home by offering direct and remote visualization of physiological parameters and enabling audio-video communication between discharged patients and healthcare staff. Integration with Margherita 3 electronic clinical records (MEDIACLINICS S.r.l.), designed specifically for neonatal intensive care units, is also implemented. The wearable devices used in Pilot A are employed in Pilot B for remote monitoring purposes.

Device usage varies according to each pilot's objectives. In Pilot A, during the first two hours in the delivery room, newborns delivered spontaneously and eligible for Skin-to-Skin Contact (SSC) will wear the ComfTech HOWDY BABY device. This early monitoring is essential for identifying potential Sudden Unexpected Perinatal Collapse (SUPC) risks from birth. Subsequently, during the first 48-72 hours in the rooming-in ward, all newborns, irrespective of delivery method, will be monitored using both the ComfTech HOWDY BABY device and the OXY-10 pulse oximeter. Intermittent monitoring will be conducted beyond the initial two hours to detect early signs of SUPC, such as bradycardia, bradypnea, and apnea, outside standard clinical monitoring periods. Healthcare staff will utilize the HOWDY App for data visualization and report generation, contributing to the refinement of rooming-in protocols.

In Pilot B, the primary aim is to facilitate the NICU-to-home transition. The MHP platform, integrated with electronic medical health records, plays a crucial role in this process. Families will access a mobile application to view and transmit their child's physiological data. Tele-visits will be conducted for 15 days following NICU discharge, enabling real-time audio-video interactions with healthcare staff. This continuous remote support aids in the prompt identification of clinical issues and offers ongoing assistance to parents.

In terms of data management, the ComfTech HOWDY BABY devices and the OXY-10 pulse oximeter automatically collect data on heart rate, respiratory rate, and oxygen saturation. This data is displayed via the HOWDY App for Pilot A. For Pilot B, the MHP platform collects physiological data and manages audio-video communications. Data collected from the HOWDY App in Pilot A and from the MHP platform in Pilot B will be exported to Excel for statistical analysis.

### 3.3.5. Evaluation metrics and statistical analysis

The characteristics of the study population are described using absolute frequencies and percentages for categorical variables, with mean  $\pm$  standard deviation or median and interquartile range applied for continuous variables. Statistical tests are selected based on data distribution.

In Pilot A, the distribution of SUPC prodromal events (including total number, type, period, and birth type) detected by the device outside standard clinical monitoring is analyzed. An event is classified as "outside" if it is detected more than 5 minutes from scheduled midwife monitoring. The analysis includes quantification of the lead time provided by the device in detecting these events. Device tolerability is evaluated through an 8-item questionnaire, targeting  $\geq 75\%$  "NO" responses for discomfort-related items (1-4, 7, 8) and "YES" responses for comfort-related items (5, 6). Correlation analyses are performed between these item responses and overall parent satisfaction, assessed through questions 9 and 10. Additional secondary objectives involve comparing exclusive breastfeeding rates, SUPC event occurrences, and associated mortality/morbidity between monitored and unmonitored groups using contingency tables.

For Pilot B, a historical baseline (2023-2024) was established prior to the initiation of the study. The baseline median (Interquartile Range - IQR) length of hospitalization was identified as 44.5 (31-68.3) days for 72 very preterm infants (gestational age  $< 32$  weeks), and 6 (3-16) days for 164 moderately (gestational age  $\geq 32$  and  $< 34$  weeks) and late preterm infants (gestational age  $\geq 34$  and  $< 37$  weeks). A baseline readmission rate of 19.4% was observed for very preterm infants. Clinical neonatal outcomes are assessed and compared against this historical cohort, focusing on primary outcomes such as length of stay, early readmission rates, and emergency care service utilization. Parent satisfaction, considered a secondary outcome, is evaluated through parent-reported experience questionnaires and interviews.

### 3.3.6. Results and discussion

No results have been produced at this time. The ethics committee protocol for Pilot A has been finalized, and resolution of the conditional approval is currently pending. Device procurement procedures are being initiated. For Pilot B, the integration program has been defined, encompassing family-centered care, telemonitoring, and real-time audio-video communication after NICU discharge, alongside the analysis of the historical baseline. Preparation of the study protocol for submission to Pilot B's ethics committee is currently underway.

### 3.3.7. Next steps

The plan delineates key activities and milestones scheduled between Month 30 and Month 48 of the project. The objective is to initiate the first clinical trial, Pilot A, by September 2025 at the latest. Given the historical data of approximately 2000 births per year at our center with the study's inclusion characteristics, and an estimated 50% study adherence rate, we anticipate enrolling about 1000 patients. This trial will implement a comprehensive protocol specifically designed for term infants. The goal is to secure Ethics Committee (EC) approval for the Pilot B clinical trial by September, enabling progression with the study focusing on preterm infants.

Regarding dissemination, an abstract has been submitted to the Joint European Neonatal Societies (JENS 2025) conference. This abstract highlights an integrated model of Family-Centered Care, combined with post-discharge telemonitoring and televisits for NICU graduates, scheduled for presentation in Autumn 2025. Additionally, an abstract is planned for submission to the Congress of the Italian Society of Neonatology to present further insights from the project. Furthermore, at least one article will be submitted to an internationally indexed journal, focusing on the monitoring of SUPC prodromal events to disseminate research findings to the broader scientific community.

### 3.4.Task 5.4 – Therapy optimization and prevention of adverse events in diabetes management (responsible: UNIROMA2, co-responsible: UNIPD)

#### 3.4.1. Overview

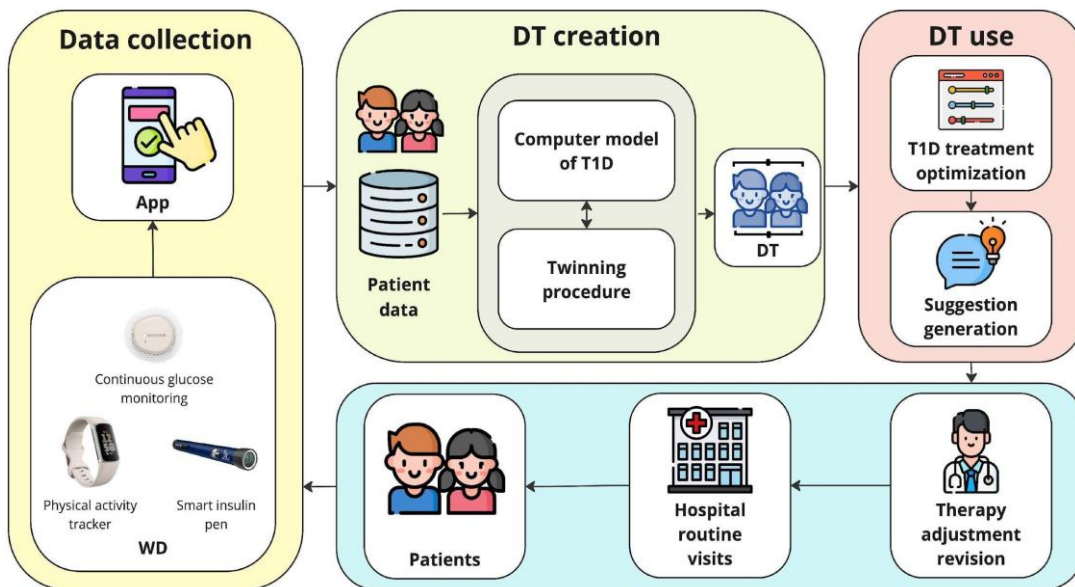


Figure 2.4.1: Graphical abstract of Task 5.4.

The project aims to create a new decision support system for clinicians managing pediatric type 1 diabetes (T1D) patients. Utilizing a dedicated mobile app, the system will gather data from various sensors, including continuous glucose monitors (CGM), smart insulin pens, and physical activity trackers, to construct patients' digital twins. This will enable, through multiple ad-hoc simulations, the optimization of basal-bolus insulin every two weeks. The optimization results will then offer clinicians advanced data analysis tools and insulin therapy recommendations.

Clinicians will access the system through a dedicated web interface, enabling them to visualize and analyse patient data using additional features such as automatic pattern recognition, customizable and comparative data visualizations, and automatically generated data summaries, which will enable them to better identify areas of therapeutic improvement.

The primary outcomes of the project are the acceptance rate of therapy modifications suggestions, system safety, and enhanced clinical workflow during patients visits using the interface.

### 3.4.2. Data used for preliminary assessment and validation

For the preliminary assessment and validation of the models, anonymized data from the Tidepool Big Data Donation Project (<https://www.tidepool.org/bigdata>), an opt-in initiative where individuals with Type 1 Diabetes (T1D) voluntarily donate their data for research purposes, were utilized. The dataset is de-identified by Tidepool prior to being licensed to academic institutions and companies for analysis. It comprises data from approximately 300 individuals with T1D, encompassing several months of continuous glucose monitoring (CGM) readings, insulin dosage (both bolus and basal), and detailed meal records.

For validation purposes, focus was placed on data from five pediatric patients aged 4 to 12 years. The dataset underwent preprocessing through a multi-step pipeline. Initially, the data were retimed on a homogeneous 5-minute time grid, consistent with the CGM sampling frequency. Subsequently, as commonly practiced in the literature, CGM data gaps of 30 minutes or less were imputed using linear interpolation. Finally, for each patient, the resulting dataset was partitioned into daily traces, yielding a total of 724 high-quality daily CGM traces (~145 per patient over a 24-week period) selected for analysis. This selection ensures the provision of reliable data for the development of algorithmic strategies.

### 3.4.3. Devices employed in the study

DEVICE #1	
<b>Model</b>	Fitbit Charge 6
<b>Vendor</b>	Fitbit
<b>What is it measuring? (also provide sampling time)</b>	Heart rate (1 sample per second), Sleep, Steps (cumulative measure every 30s), Physical activity
<b>Total number of devices used in the study</b>	60
<b>Describe how it integrates in your system</b>	It allows to collect physical activity data, useful for both models and clinicians during visits

DEVICE #2	
<b>Model</b>	G7
<b>Vendor</b>	Dexcom
<b>What is it measuring? (also provide sampling time)</b>	Glucose concentration (one sample every 5 minutes)
<b>Total number of devices used in the study</b>	1440 (one per patient every 10 days)

<b>Describe how it integrates in your system</b>	It allows to continuously collect blood glucose concentration, essential for T1D management
--	---

<b>DEVICE #3</b>	
<b>Model</b>	NovoPen Echo Plus
<b>Vendor</b>	NovoNordisk
<b>What is it measuring? (also provide sampling time)</b>	Injected insulin dose
<b>Total number of devices used in the study</b>	80
<b>Describe how it integrates in your system</b>	It allows to log and store insulin doses injected by the patient

#### 3.4.4. Developed models/algorithm/platforms

The developed decision support system for clinicians managing pediatric T1D patients consists of three main components. The first component is a mobile application for collecting patient data. The second component, which represents the algorithmic core of the system, is a module that leverages data collected from the mobile app to generate T1D basal-bolus insulin therapy optimization recommendations. Finally, the third component consists of a web interface aimed at presenting such recommendations to clinicians and providing advanced and customized data analysis tools aimed at optimizing routinary patients' visits. Each component is described below in detail.

##### ➤ Mobile application for patient data collection

The mobile application is the data collection hub for the patient. It allows patients to interact with the system, collecting insulin data from smart insulin pens, glucose data from CGM sensors, and physical activity data from physical activity trackers. Its features and design have been developed following the feedback collected during two dedicated in-person co-design workshops, aimed at understanding and aligning to patients' therapy management habits, thus reducing the data collection burden. Furthermore, the application is equipped with interactive gamification features, which will improve patients' adherence to data collection during the upcoming clinical trial.

##### ➤ T1D basal-bolus insulin therapy optimization module

The system aims to provide clinicians with biweekly insulin therapy recommendations tailored to the specific needs of each pediatric patient, as shown in Figure 1.

After collecting data over a two-week period and generating daily digital twins, the optimization process starts. The system minimizes the Glycemic Risk Index (GRI) over the 14-day window by leveraging digital twin simulations created using the ReplayBG tool (Capon G, Vettoretti M, Sparacino G, Favero SD, Facchinetti A. "ReplayBG: A Digital Twin-Based Methodology to Identify a Personalized Model From Type 1 Diabetes Data and Simulate Glucose Concentrations to Assess Alternative Therapies. *IEEE Trans Biomed Eng.* 2023;70(11):3227-3238.) Specifically, the system explores various combinations of modulation parameters for basal insulin ( $k_{bi}$ ) and the three main meal-time boluses ( $k_b$  for breakfast,  $k_l$  for lunch, and  $k_d$  for dinner), for simplicity represented as a vector  $k$  in Figure 1. For each combination, it simulates daily glycemic responses and computes the corresponding GRI. The optimization algorithm iteratively tests different parameter sets until it identifies the combination that yields the lowest overall Glicemia Risk Index (GRI) across the biweekly period. This optimal set of modulation parameters  $k^*$  is then presented to clinicians as a therapy adjustment suggestion.

Inputs include patient continuous glucose readings, insulin dosing history (both basal and bolus), and meal carbohydrate intake. The output is a set of optimal modulation parameters for basal insulin and meal-time boluses, intended to be applied during the subsequent biweekly period.

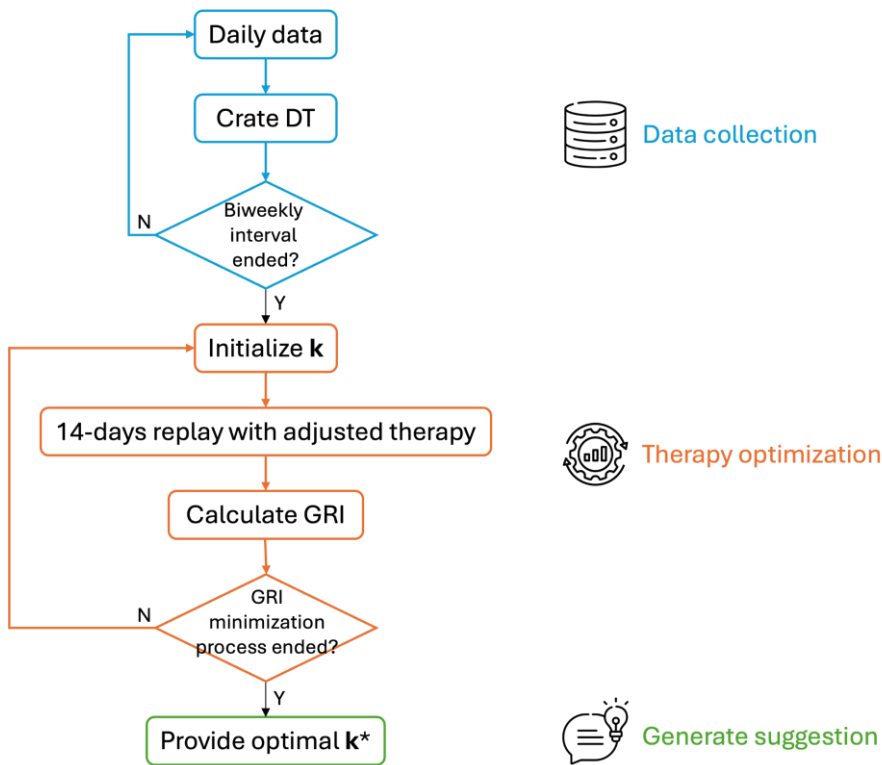


Figure 2.4.2: Flowchart showing the therapy optimization process.

➤ Web interface

The web interface aims to provide clinicians with a user-friendly platform for monitoring and analyzing glucose data and evaluating algorithm-generated recommendations. It presents the results of the T1D basal-bolus therapy optimization module and analyses collected patient data. Particularly, the user interface for patients' data visualization has been designed in collaboration with clinicians during two dedicated co-design workshops, ensuring that the developed features are aligned with their requirements. Figure 2 shows the mockup final design. Each block is linked to a specific step performed during the visit, starting from Block A, where the dashboard shows relevant glucose metrics and patient information. Blocks C and D contain the main analysis visualizations, while Block B is reserved for the narrative generated summary. Block E is dedicated to visualizing the therapeutic suggestions obtained by the module described above.

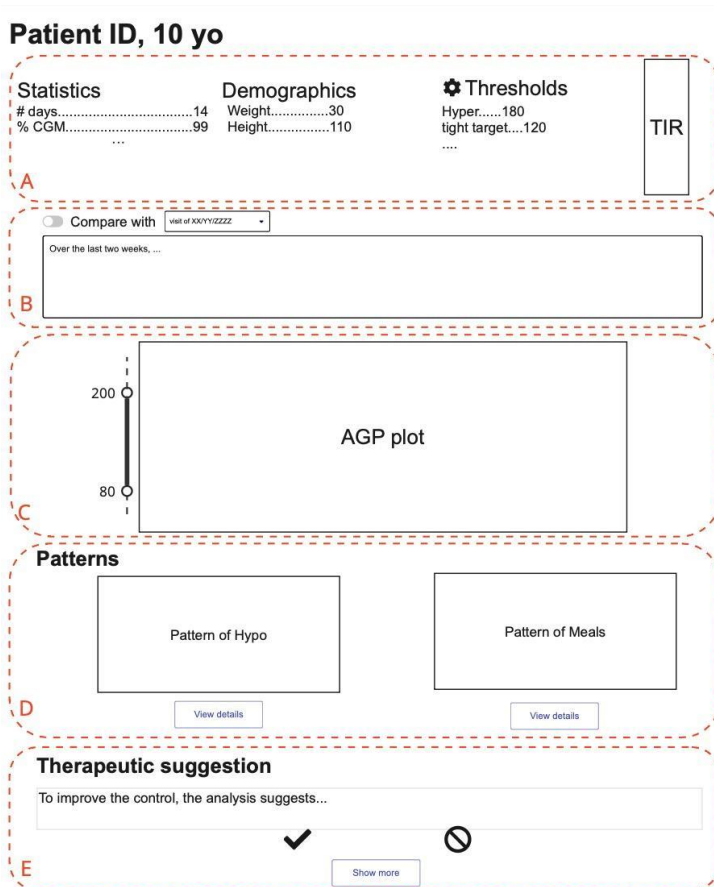


Figure 2.4.3: Mockup of the data analysis web interface.

➤ Secure data management and interoperability aspects

All data collected and managed by the platform are stored in an encrypted format (AES-256). Communication between platform components occurs via end-to-end encrypted channels and employs a role-based authentication system, ensuring that only authorized users can access sensitive data.

The system does not connect to existing EHR systems since it is still experimental. All data is collected and managed internally by the platform to facilitate necessary data processing and usage across all components (i.e., the algorithmic core and advanced data visualization on the website). Access to raw data is granted upon request through dedicated secure web APIs, allowing authorized researchers to download fully anonymized data in a structured and documented format. All access is logged for auditing purposes. While the system does not strictly adhere to health data format standards, it utilizes an ad-hoc data structure that ensures minimal data storage in accordance with GDPR regulations.

To ensure data consistency and mitigate security risks, periodic backups of patient data will be performed and securely stored on separate disks. All platform features will undergo regular sanity checks to validate their correct behavior and status.

The foundational component of the algorithmic core (ReplayBG) is open source and can be accessed at the following link ([https://github.com/gcappon/py\\_replay\\_bg.git](https://github.com/gcappon/py_replay_bg.git)). Currently, to maintain confidentiality, all platform code is stored in versioned private repositories. In later stages of the project, certain parts of the code may be open-sourced to promote reproducibility.

### 3.4.5. Evaluation metrics and statistical analysis

Each component of the system has been evaluated using ad-hoc strategies defined below.

#### ➤ Mobile application for patient data collection

The mobile application prototype has been evaluated in two dedicated codesign sessions with pediatric patients and their caregivers. During these sessions, the evaluation was limited to a qualitative assessment of the application design and features, receiving verbal feedback and suggestions.

#### ➤ T1D basal-bolus insulin therapy optimization module

A 24-week in silico clinical trial was designed to evaluate the optimization framework, simulating biweekly therapy adjustments by clinicians.

Glucose control achieved through optimized insulin modulation was assessed over each 14-day interval using clinically accepted metrics: time-in-range (TIR, blood glucose within 70–180 mg/dL), time-above-range (TAR), time-below-range (TBR), coefficient of variation (CV), and glycemia risk index (GRI).

To evaluate longitudinal performance, the median daily value of each metric per biweekly interval across subjects was computed, followed by the fitting of a robust linear model with a bisquare weighting function. The slope of each fitted model served to assess the direction and strength of trends over the 12 intervals.

#### ➤ Web interface

To assess the web interface's usability, five clinicians were asked to complete a series of predefined tasks designed to explore all the relevant features of the interface and reproduce a routine visit. Each task was rated in a 5 Likert scale and the overall experience was evaluated by the System Usability Scale (SUS) questionnaire, which evaluates in a scale from 0 to 100 the overall system ease of use.

### 3.4.6. Results and discussion

#### ➤ Mobile application for patient data collection

The first prototype received positive feedback during the codesign sessions. Feedback on current features of the prototype included suggestions for possible improvements. Additionally, common interactions with commercial diabetes management mobile applications were identified to align the new application with established user habits, thereby reducing the usage burden. Potentially effective gamification strategies were also identified to enhance adherence to clinical protocols in data collection over extended periods, as supported by literature in multiple use cases.

#### ➤ T1D basal-bolus insulin therapy optimization module

During the 24-weeks in silico clinical trial, results demonstrated the progressive impact of adaptive insulin therapy adjustments on glycemic control. TIR showed an upward trend across biweekly intervals, reaching 72.12% in weeks 23–24 and meeting the clinical target of  $\geq 70\%$ . TAR and TBR lower throughout the trial, indicating fewer hyperglycemic and hypoglycemic episodes. The GRI also shows a reduction across the simulation period, with notable improvements emerging after the initial optimization cycles. This decline reflects the effectiveness of the optimization strategy in reducing glycemic risk over time.

As exemplified for TIR in Figure 3, analyzing the longitudinal trends in TIR, TAR, TBR, GRI, and CV, the slopes of the respective linear models suggest incremental improvements in glycemic control over the trial period. Particularly, on average, every two weeks, TIR increased by +0.63%, indicating that patients are spending more time within the target glucose range, a key indicator of effective diabetes management; while TAR and TBR decreased by -0.18% and -0.25%, respectively, suggesting reductions in both hyperglycemic and hypoglycemic episodes, which contribute to a more stable glycemic profile. GRI declined by -0.64, and coefficient of variation (CV) showed a slight reduction of -0.16, suggesting improved stability in glucose levels.

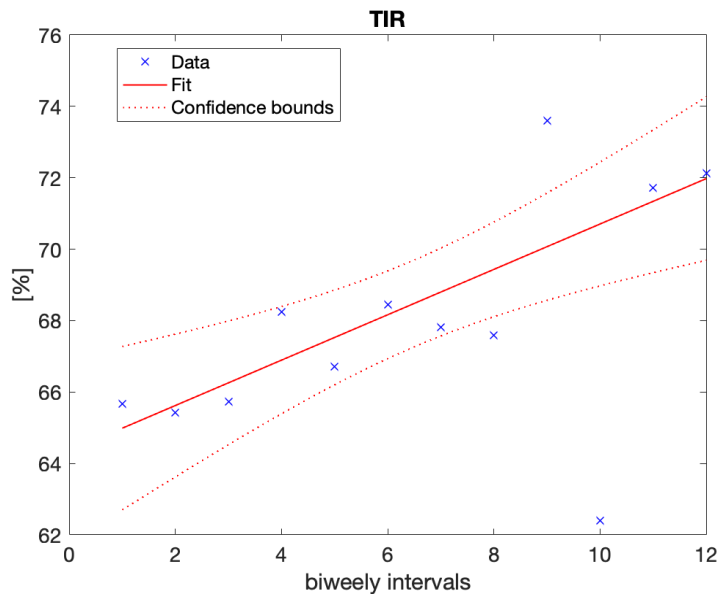


Figure 2.4.4: Example of the fitted linear model for TIR metric.

➤ Web interface

The interface received a median SUS score of 82.5, indicating good usability. Clinicians found features like customizable parameters and enhanced analysis visualizations beneficial for their routine visit workflow. Of note, questions related to ease of use and ease of learning the system obtained high marks among all responders. Oral feedbacks were positive and minor graphic suggestions were collected and will be implemented in the future versions of the system, to further improve its usability and adherence to the clinical requirements.

3.4.7. Next steps

During the period spanning months 30 to 48 (M30–M48), the complete system is scheduled to be finalized, deployed, and tested. By the end of M32, all system components are expected to be refined based on feedback gathered during preceding months and fully deployed within the production server infrastructure. This deployment will facilitate the initiation of the clinical trial, which is scheduled to commence between M33 and M35. The trial will extend over six months, during which continuous monitoring and fine-tuning of the system will be conducted, emerging issues will be addressed, and preparations for trial data analysis, slated to begin in M40, will be undertaken. Of note, the goal is to secure Ethics Committee approval by M33 and no later than M35, enabling the timing conduction of the clinical trial.

Project outcomes are intended to be disseminated at prominent international conferences in the field of diabetes care, such as the European Association for the Study of Diabetes (EASD) annual meeting and the International Society for Pediatric and Adolescent Diabetes (ISPAD) annual conference, as well as in technology-oriented forums, including the Advanced Technologies and Treatments for Diabetes (ATTD) and the Diabetes Technology Meeting (DTM). Furthermore, at least three articles are planned for submission to internationally indexed journals: one focusing on the optimization module for T1D basal-bolus insulin therapy, another detailing the design, development, and evaluation of the web interface, and a third presenting the results of the forthcoming clinical trial.

In terms of exploitation, engagement with the ReactorPro initiative is ongoing, aimed at supporting the long-term sustainability and potential future commercialization of the platform as a software as a medical device (SaMD).

### 3.5.Task 5.5 – Non-medical wearable devices for monitoring caloric intake (responsible: UNIPD)

#### 3.5.1. Overview

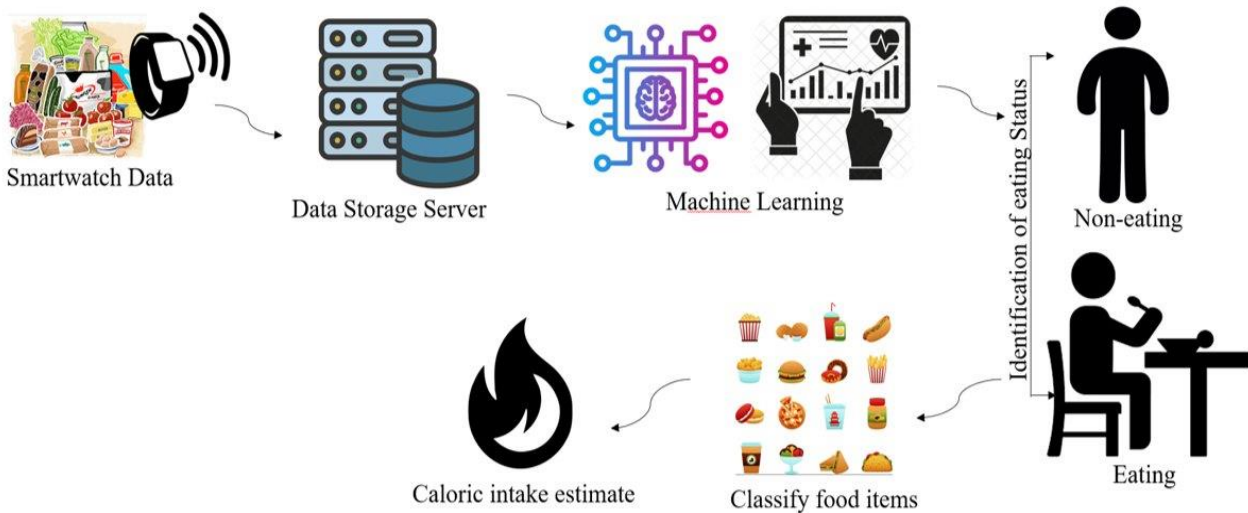


Figure 2.5.1: Graphical abstract of Task 5.5.

Monitoring lifestyle, and in particular calorie intake and energy expenditure, is currently a crucial challenge for maintaining individual and public health. Traditional methods (i.e., questionnaires) do not reach an acceptable level of accuracy, having an estimated error of 25% on the estimate of daily caloric intake. This study aims to develop a precise and stable algorithm for estimating caloric intake using machine learning methodology. The study will capture categorized individual eating activities firstly in target populations (i.e., people with special diet requirements) for algorithm development, then in healthy (i.e., with no major pathologies or particular conditions that limit movements or require specific diets) Italian people. Raw data from wearable devices will be analysed through an up-to-date machine learning approach to accurately predict the caloric contribution of individual eating activities.

#### 3.5.2. Data used for preliminary assessment and validation

The data used to assess and validate the models were obtained from 20 healthy adult participants (9 women and 11 men, aged 20–30 years) recruited for the NOTION study, which explored the use of commercial smartwatches (Garmin Fenix 5) to detect eating behaviors using motion sensor data. The dataset, comprising 278,260 timepoints, was collected in a semi-naturalistic university cafeteria setting during four meal occasions (breakfast, lunch, snack, and dinner) and involved the synchronized recording of accelerometer data with

video annotations of eating behavior. Preprocessing included converting FIT-format data to CSV, feature extraction, and generation of windowed time-series summaries (mean, median, SD, min, max, IQR, slopes) for each kinetic variable. Each window was labeled as "eating" or "non-eating" based on majority vote from manually annotated frames. To ensure data representativeness, meals were diverse and included 13 food items, while participants were allowed to move freely in a naturalistic dining environment. The dataset was randomly split into training (80%) and test (20%) sets with stratified class distribution, and 10-fold cross-validation was applied on the training set to prevent overfitting and ensure robust model evaluation. Techniques like dummy encoding, zero-variance filtering, and normalization were applied, and multiple models were compared using balanced accuracy and ROC metrics. Independence between training and test sets was maintained throughout.

### 3.5.3. Devices employed in the study

<b>DEVICE#1</b>	
<b>Model</b>	Garmin Fenix 5 Smartwatch
<b>What is it measuring? What is the role of the device in the study?</b>	The Garmin Fenix 5 is a commercial smartwatch equipped with inertial sensors (accelerometers and gyroscopes). It was used to capture motion data from both wrists of each participant, including acceleration along the x, y, and z axes, pitch (tilt), roll (rotation), power (magnitude of acceleration), and total energy (gravity-corrected power). The primary role of the device was to non-invasively monitor and record wrist movements during eating episodes to enable the detection and classification of eating behavior using machine learning algorithms.
<b>Total number of devices used in the study</b>	2 devices

<b>DEVICE#2</b>	
<b>Model</b>	FitBit Sense
<b>What is it measuring? What is the role of the device in the study?</b>	The Fitbit Sense is a commercially available smartwatch equipped with multiple sensors, including a 3-axis accelerometer and gyroscope. In this study, it was employed to non-invasively collect motion data from the participant's wrist during daily activities. The recorded features included triaxial acceleration, angular velocity (pitch, roll, yaw), signal magnitude vector (SMV), and derived energy

	metrics. These data streams were used to characterize wrist movement patterns during specific behaviors, particularly food intake, enabling downstream analysis and classification through machine learning approaches.
<b>Total number of devices used in the study</b>	6 devices

<b>DEVICE#3</b>	
<b>Model</b>	Apple Watch SE
<b>What is it measuring? What is the role of the device in the study?</b>	The Apple Watch SE is a consumer-grade smartwatch featuring integrated inertial sensors, including a 3-axis accelerometer and gyroscope. In this study, it was utilized to passively record wrist motion data during naturalistic behavior. Captured signals encompassed triaxial acceleration and angular velocity, from which features such as pitch, roll, yaw, signal magnitude vector, and kinetic energy were derived. The device enabled continuous, unobtrusive monitoring of upper limb movements, supporting subsequent identification and classification of behavioral episodes, such as eating, via machine learning methodologies
<b>Total number of devices used in the study</b>	2 devices + participants' owned devices

<b>DEVICE#4</b>	
<b>Model</b>	Garmin Vivoactive 3, 4, and 5
<b>What is it measuring? What is the role of the device in the study?</b>	The Garmin Vivoactive 3, 4, and 5 are multisport smartwatches equipped with embedded inertial measurement units, including 3-axis accelerometers and gyroscopes. Across all three models, these sensors provide continuous acquisition of wrist motion data, including triaxial acceleration and angular velocity. For this study, the devices were employed to unobtrusively monitor wrist movements in free-living conditions, enabling the extraction of derived features such as pitch, roll, yaw, signal magnitude vector, and energy metrics. The recorded data supported the identification and classification of specific behaviors, particularly eating episodes, through the application of machine learning algorithms.

<b>Total number of devices used in the study</b>	5 devices + participants' owned devices
--	---

### 3.5.4. Developed models/algorithm/platforms

#### ➤ Web server

The UBEP Wearables Dashboard is an advanced platform developed in collaboration with Sinotropy.ai (Milan, Italy). Designed to manage wearable devices across multiple research studies, the platform enables seamless reassignment of devices to different studies as needed. It is device-agnostic, allowing the integration, analysis, and visualization of both raw and aggregated data from various sensors. The system operates through API-based interactions, directly querying web APIs or utilizing an intermediate server for low-level device APIs. The interface supports real-time data synchronization and customizable analytics pipelines, ensuring flexible and scalable data-driven assessments for diverse research applications.

#### ➤ RedCap/MyCap Food Questionnaire

To validate the data collected from these wearable devices, a food frequency questionnaire (FFQ) have been developed in collaboration with clinical nutritionists from UNIBA, specifically tailored for patients with polycystic kidney disease adhering to a ketogenic diet (Task 5.2). The FFQ was implemented via RedCap/MyCap, allowing participants to digitally record comprehensive information on their dietary intake. In parallel, participants maintained a daily food diary in the same app/platform to document all meals consumed throughout the day in greater detail. App-based reminders were employed to enhance compliance and completeness. Additionally, participants were instructed to take photographs of their meals before and after consumption and upload them to the platform, enabling the use of visual data to further validate intake through image-based models.

#### ➤ Objective

To develop an end-to-end system using smartwatch sensor data to automatically classify time segments as eating or non-eating.

#### ➤ Key Functioning

Two Garmin Fenix 5 smartwatches, worn on both wrists, recorded accelerometer and gyroscope data at 5 Hz during meals under semi-naturalistic conditions. Video recordings with hand-clap signals allowed synchronization for manual labeling. Raw data in FIT format were converted to CSV and processed in R.

Motion data were segmented into 1–5 second windows and summarized into Information Units (IUs) using statistical (mean, median, SD, etc.) and dynamic (slope) features. Preprocessing included dummy encoding and normalization.

These structured, labeled IUs served as inputs for various machine learning models, including interpretable (Decision Tree, Random Forest), black-box (Neural Network, XGBoost, GBM), generative (Naive Bayes, Linear Discriminant Analysis - LDA, BART), and instance-based (K-Nearest Neighbors - KNN) classifiers. Models were trained on 80% of the data and validated on 20%, with 10-fold cross-validation. Random Forest and Decision Tree models performed best, achieving high accuracy (up to 0.837) and strong ROC scores (up to 0.958).

➤ Input/Output of the system

Input of the system are sensor data (FIT format), annotated time labels, feature-engineered windows while the output is the binary classification of each time window as eating or non-eating, with associated performance metrics.

### 3.5.5. Evaluation metrics and statistical analysis

To assess the performance of machine learning models in classifying eating and non-eating episodes using smartwatch sensor data, several widely accepted evaluation metrics and validation techniques were employed. These include:

- **Balanced Accuracy:** the primary evaluation metric, as it accounts for class imbalance by averaging sensitivity (true positive rate) and specificity (true negative rate). Balanced accuracy is widely used in activity recognition and biomedical applications, particularly where the dataset is skewed-as is typical in eating vs. non-eating detection tasks.
- **Sensitivity (Recall):** Measures the proportion of correctly identified eating episodes among all true eating episodes.
- **Specificity:** Measures the proportion of correctly identified non-eating episodes.
- **Overall Accuracy:** Proportion of all correctly classified instances.
- **Area Under the Receiver Operating Characteristic Curve (ROC-AUC):** This reflects the model's ability to discriminate between the eating and non-eating classes across all thresholds. AUC values close to 1.0 indicate strong classification performance.

These metrics are standard and extensively reported in the literature related to human activity recognition using wearable sensors, ensuring comparability with prior work.

## Validation Techniques

To ensure robust evaluation of model performance and minimize overfitting the dataset was randomly split into training (80%) and test (20%) sets while preserving class distribution (stratified split). Furthermore, a 10-fold cross-validation strategy was applied on the training set to improve the reliability of the performance estimates and prevents dependence on a particular random split.

To ensure the integrity of model validation and prevent data leakage several preprocessing steps including dummy encoding, normalization, and removal of zero-variance predictors, were applied within the training set only using the caret package in R, ensuring that test data remained unseen during these steps. Feature selection and parameter tuning (e.g., choice of optimal variable combinations or model hyperparameters) were conducted strictly using cross-validation within the training data and the final test set was used only once after all training and tuning were completed, preserving its independence for unbiased performance evaluation.

### 3.5.6. Results and discussion

The evaluation of machine learning models for detecting eating behavior (i.e., eating vs. non-eating data segments) using smartwatch sensor data yielded highly encouraging results. Among the tested classifiers, Random Forest emerged as the top-performing model, achieving a balanced accuracy of 0.837 and an ROC-AUC of 0.958 on the test dataset. This indicates excellent discriminatory power and robust generalization to unseen data. Decision Tree and K-Nearest Neighbors (KNN) models also performed well, with balanced accuracies of 0.826 and 0.808, respectively, and high sensitivity scores, highlighting their effectiveness in correctly identifying eating episodes. These results are consistent with existing literature, where ensemble and tree-based models consistently outperform others in human activity recognition tasks using inertial sensor data. The use of balanced accuracy as the primary metric is particularly relevant, given the inherent class imbalance between eating and non-eating episodes, thereby reinforcing confidence in the reliability of the findings.

Despite promising results, the current system is not yet ready for full deployment. The models were validated under semi-naturalistic conditions with a limited sample of 20 participants, which may not capture the diversity of eating behaviors across different populations, cultures, and settings. Key challenges remain,

including misclassification due to non-eating hand movements, the need for scalable annotation methods, and integration with mobile platforms.

Future development efforts will focus on expanding the participant base, collecting real-world free-living data, and designing a user-friendly interface for clinical or consumer use. These limitations are typical in early-stage wearable monitoring systems and can be addressed through iterative refinement. The results confirm the technical feasibility of using commercial smartwatches with machine learning for dietary monitoring. Ongoing efforts are directed towards these improvements, with system deployment planned for September.

### 3.5.7. Next steps

Over the next 18 months (M30–M48), the project will focus on finalizing the system for detecting eating behavior, developing a new algorithm for classifying food items, and preparing the platform for real-world deployment. Key activities will include expanding data collection with a larger and more diverse participant pool in free-living settings, refining the preprocessing and classification pipeline, and integrating the current eating detection model with a new module designed to identify specific food items based on sensor patterns. A lightweight, real-time implementation will be developed for mobile or wearable platforms, incorporating secure data handling, user authentication, and an intuitive interface. Usability testing will be conducted with a pilot group to assess the system's performance, reliability, and user experience. The final phase will involve optimizing the system, completing technical documentation, and packaging the full software pipeline, encompassing both eating detection and food classification components, for reproducibility, dissemination, and potential deployment.

### 3.6.Task 5.6 – Prevention/mitigation of frailty in the continuum of care framework (responsible: UNIROMA2)

#### 3.6.1. Overview

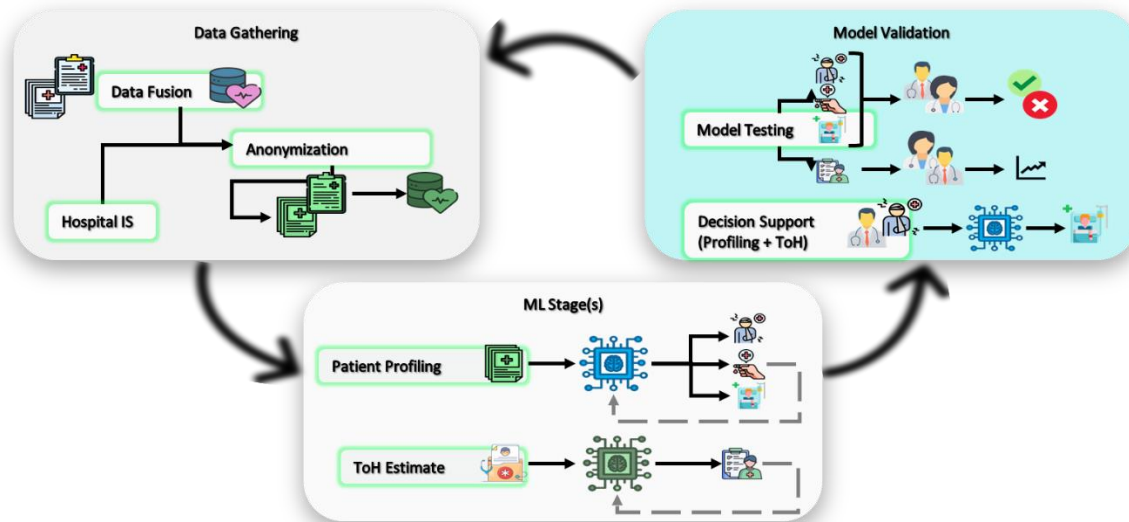


Figure 2.6.1: Graphical abstract of Task 5.6.

The project aims to lay the groundwork for developing a model of integrated healthcare and social care, supported by telemedicine and AI.

The development of this model will take into account the guidelines provided in the Decalogue for the implementation of national healthcare services through Artificial Intelligence systems.

The overall objective is to create an algorithm based on Machine Learning (ML) techniques to predict the length of hospital stay for patients: MARIO - Modello di Analisi del Rischio di uso Improprio dei servizi Ospedalieri (RAMIS - Risk Analysis Model of Inappropriate use of hospital Services). This model will provide useful insights to help prevent Over-Threshold (OT) hospitalizations and reduce the percentage of Frequent Users (FU) of healthcare services.

The research project involves the development of the predictive system through a retrospective cohort study, based on routinely collected information by hospital staff and provided to the Health Directorate of the Tor Vergata University Hospital (PTV), which is involved in the project.

The data used will be obtained from Hospital Discharge Records and from the Emergency Information Management System.

### 3.6.1. Data used for preliminary assessment and validation

This research project will be conducted through a retrospective cohort study involving two groups, for which a single predictive model is expected to be developed. The model will rely on information routinely collected by hospital staff and made available to the project team by the Health Management Unit of the Tor Vergata University Hospital (PTV).

The predictive model will be built using data from hospital admissions recorded in 2022, 2023, and 2024. The data will be sourced from Hospital Discharge Records and the Emergency Information Management System (GIPSE).

A sample of approximately 10,000 hospital admissions will be selected for model development. Aggregated and anonymized data extracted from the Hospital Discharge Records and the GIPSE system will be used to identify Over-Threshold admissions and Frequent Users.

Sample size: The target reference population includes approximately 10,000 hospital admissions for patients aged over 65, and about 37,000 annual Emergency Room visits.

### 3.6.2. Devices employed in the study

<b>DEVICE#1</b>	
<b>Server</b>	Standard PC technology
<b>Objectives and targeted phenomena</b>	Data Gathering from the Hospital: it is the transient node from private data (Hospital records) to anonymized data.
<b>Total number of devices used in the study</b>	1

<b>DEVICE#2</b>	
<b>GPU Server</b>	Medium powered GPU hardware for supporting training and testing under neural and generative AI technologies
<b>Objectives and targeted phenomena</b>	Machine Learning over anonymized data. It is achieving optimal models for risk classes induction and estimates (regression) of ToH

<b>Total number of devices used in the study</b>	2
--	---

### 3.6.3. Developed models/algorithm/platforms

In the current development stage of the project the adopted systems are only simple Data Management systems and Machine Learning software, ranging from standard approaches such as Support Vector Machine (SVM) or Bayesian software libraries, to deep learning architectures and algorithmics. No interaction with the final users is actually needed, so that some concerns about interpretability of results, accessibility and security issue have a minor impact in this stage.

The major concern of the period has been Data Management with its aspects related to (1) deciding the operational data sources for the project work (experimenting Machine Learning and validating the resulting data-driven time of hospitalization (ToH) models), (2) studying the privacy aspects related to the targeted data, (3) designing a technology for pseudo-anonymization in coordination with the hospital referents (Data Protection Officer team) and (4) implementing the anonymization software for preparing the final data release for training and validating the project model obtained via ML.

Tools resulting from this stage are simple batch services for the medical staff, for ensuring reliable, secure and expressive anonymization able to remove personal information from the input, and release a still expressive yet traceable form of output. In other words personal information will be removed but crypted as in modern secure coding systems so that records of the same person, even not knowing (inversion) which person it is, could be recognized in the anonymized data and used for training patient descriptors that characterize several clinical events along the temporal axis.

The adopted workflow will treat data within the hospital firewall before encryption and only anonymized records will be made available for ML. The learning thus will act on impersonal data, whose details about identities cannot be reconstructed in any way.

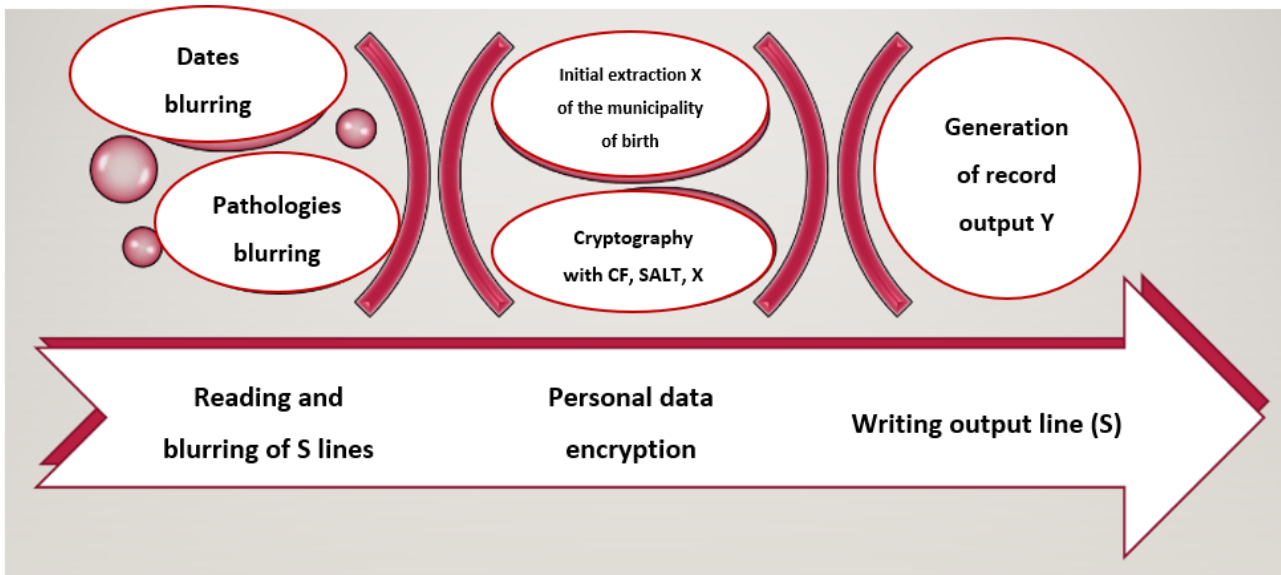


Figure 2.6.2: Main steps of the anonymization process followed in MARIO.

The described process initiates with the obfuscation of specific risk sources associated with identity recognition (Figure 2.6.2, left), progresses through the encryption stage (Figure 2.6.2, center), and concludes with the compilation of output data records (Figure 2.6.2, right).

In the initial step, blurring of dates and pathology information involves minimal manipulations designed to preserve the majority of the original data record's intended information. These modifications are crafted not to impact the quality of the Machine Learning (ML) stage while being randomized sufficiently to prevent tracing back to individual clinical events (dates) or identifying personal (potentially rare) pathologies. Special attention is given to manipulating pathology codes to ensure that even individuals with rare illnesses, who belong to small reference communities that could otherwise facilitate identity reconstruction, remain unidentifiable.

Blurring introduces slight noise to dates but primarily alters pathology codes by reducing their granularity. This is achieved by manipulating Health Level Seven (HL7) and other regional standard codes, such as removing the rightmost character of an alphanumeric code, thereby eliminating specific details. These modifications do not compromise the effectiveness of ML processes but make precise reconstruction of patient records impossible, effectively disconnecting individual identities from the blurred data.

The second stage involves encryption, where key identification fields are recoded using the SHA-256 algorithm to generate unique, non-invertible digital fingerprints, 64-character strings representing each individual's tax code (a unique fiscal identifier in Italy). Despite their uniqueness, these encrypted codes cannot be decrypted to retrieve the original personal information. The SHA-256 algorithm, commonly employed to ensure privacy in password-based communications and blockchain applications, supports this functionality. Encryption results in consistent codes across the dataset, enabling the formation of informative temporal chains of clinical events for ML studies, without any link to personal identities. Since no identifiable details remain post-encryption and the codes cannot be reversed, anonymized data records cannot be traced back to individuals.

The final stage involves compiling output records for each clinical event. This process removes irrelevant information, blurs temporal and clinical coding fields, and, following encryption, eliminates all personal references, including names, surnames, birthdates, and other demographic details such as the fiscal code.

#### 3.6.4. Evaluation metrics and statistical analysis

At this stage of the project, at least two measures are to be executed.

Firstly, the quality of the gathered data will be assessed. This involves measuring the throughput of the anonymization procedure and its reliability, defined as the probability of code compromise. This reliability is supported by studies concerning the cryptographic method employed, analogous to those used for secure password management in contemporary Operating Systems and Networking technologies.

Secondly, the focus will be on evaluating the quality of the Machine Learning techniques utilized for ToH and Risk class estimation. For ToH, classical metrics such as precision, recall, and F1-measure will be applied when ToH is treated as a discrete variable with an enumerable range (e.g., OT, not OT, likely OT). In scenarios involving numerical estimates, quantitative metrics such as Mean Squared Error or Cross-Entropy will be employed to compare the predicted ToH with the actual hospitalization time (considered as the Oracle or a posteriori Gold Standard).

For risk class estimation, Cross-entropy will be utilized. However, precise definitions remain challenging due to the absence of a specific notion of Risk class. The establishment of project-specific Risk categories will form

part of the data analysis phase, scheduled for the upcoming months. No additional evaluation metrics beyond the aforementioned standards are currently deemed necessary as project requirements.

### 3.6.5. Results and discussion

At this stage of the project, training data are not available, as the required input training data sets have not yet been assembled.

When the process is modeled as an integrated classification and regression task over time, several technologies established in similar problem domains will be explored. Neural technologies, specifically deep learning systems for time series such as Informer (Zhou H, et al. Informer: beyond efficient transformer for long sequence time-series forecasting. arXiv, preprint, 2021. <https://doi.org/10.48550/arXiv.2012.07436>) or TimeX (Zhong X, et al. XTime: A general rule-based method for time expression recognition and normalization, Knowledge-Based Systems, Volume 297, 2024. <https://doi.org/10.1016/j.knosys.2024.111921>), are anticipated to be adopted due to their demonstrated state-of-the-art performance on non-medical datasets, including transport and car traffic data, as well as financial data.

For instance, TimeX has been proposed to generate interpretable temporal patterns applied to diverse real-world time series classification tasks, such as ECG arrhythmia detection, human activity recognition, and epilepsy or EEG seizure detection. In scenarios like arrhythmia detection, ground-truth explanations for ECG, such as QRS intervals, are derived from known regions of ECG signals where arrhythmias are detectable. TimeX has been shown to identify relevant QRS intervals critical for arrhythmia diagnosis, outperforming leading baselines by 5.39% in AUPRC and 9.83% in AUR. Additionally, TimeX's explanations demonstrate superior performance in AUR, identifying broader segments of the QRS interval rather than isolated timesteps.

On the efficiency front, the project aims to deliver an accurate yet fast anonymization batch service. The current, non-final version of the anonymization software is capable of processing over 100 health records per second, as evidenced by tests conducted on separate and simplified datasets.

### 3.6.6. Next steps

The development of the proposed anonymization technology is nearing full completion, with current performance metrics confirming its high processing efficiency. At the current pace, a sufficient volume of

anonymized training data is expected to be collected by the end of June 2025, ensuring readiness for the launch of experimental evaluations, such as ToH model training and validation, in July.

By October 2025, detailed reports will be released presenting high-quality ToH estimation results and describing a workflow management system for tracking the training progress and performance of the most promising models. In parallel, the first version of the forecasting system will be delivered, intended for preliminary integration testing within hospital information processing workflows.

Also in July, the design phase of a Risk Management Class System will be initiated. This component will categorize clinical input data into risk classes based on large-scale data analysis, with clinical validation. The design phase is scheduled to conclude by the end of September 2025, paving the way for experimental implementation and testing of automated risk classification methods starting in October 2025.

## 4. Conclusion and next phase

This deliverable highlights the progress achieved by WP5. All six pilot studies have shown important advancements regarding models preliminary assessment and validation, with some of them that already moved from the design phase into real-world implementation and early validation. Each pilot study has developed tailored digital tools adapted to specific clinical contexts, ranging from wearable-based platforms for chronic disease management to mobile health solutions for pediatric care. The research design integrated various data sources, including clinical records, imaging, genetic profiles, and wearable sensor outputs. Collection of structured datasets essential for evaluating the performance of the systems under investigation has begun.

Initial deployments have demonstrated both the technical feasibility and clinical relevance of the developed tools. Early user feedback, especially from healthcare professionals involved in the pilot studies, has already informed improvements to system usability, report generation, and alert mechanisms. In some cases, the platforms have started to show potential benefits for enhancing care coordination, patient engagement, and early detection of clinical deterioration.

Despite the encouraging progress achieved across the six pilot studies in WP5, several limitations currently constrain the full implementation and validation of the proposed solutions. In Task 5.1, while the clinician-facing digital tool is operational and early feedback has been positive, the patient-facing mobile application is still under development. Full integration between components is ongoing, and the predictive modeling functionality has yet to undergo clinical validation. In Task 5.2, although the technological infrastructure and study design are well defined, validation of the developed models is contingent on the completion of patient recruitment, which has only recently started. Task 5.3 and Task 5.4 are both awaiting final approval from the respective ethics committees. As a result, patient enrollment haven't commenced yet in pilots connected to these tasks, delaying the final validation of the developed tools. Task 5.5 has demonstrated promising preliminary results under controlled conditions, but the system still lacks the readiness required for large-scale deployment and real-world testing. Finally, Task 5.6 is focused on retrospective data analysis, and the anonymized datasets needed for training predictive models are still being collected and processed. All the one issue listed so far in this paragraph were expected at this stage of development and will be systematically addressed during the next project phase (M30–M48), as pilot activities move toward full deployment, validation, and clinical impact assessment.

Looking ahead to the next phase of the project, which spans months 30 to 48, efforts will focus on completing patient recruitment and expanding tool deployment to additional clinical centers. As longitudinal data accumulates, WP partners will concentrate on activating and evaluating the predictive capabilities of the systems under development. These evaluations will rely on rigorous statistical methodologies, including cross-validation and time-to-event analysis, to ensure robustness and generalizability. At the same time, refinements to both back-end algorithms and user-facing components will continue based on real-world usage and clinical feedback. Specific attention will be given to assessing interoperability with existing clinical infrastructures and exploring the regulatory pathways required for medical certification. Dissemination of results through scientific publications, conference presentations, and stakeholder engagement will also play a central role in this next phase.

Ultimately, the work done to date lays a strong foundation for demonstrating the real-world impact of ICT-enabled tools on patient outcomes and health system performance. The upcoming validation phase will be crucial for confirming the added value of these interventions and establishing their readiness for broader adoption.