

Article

Machine Learning-Enhanced Architecture Model for Integrated and FHIR-Based Health Data

Nadia Brancati ^{1,†} , Teresa Conte ^{1,*,†} , Simona De Pietro ^{2,3,†}, Martina Russo ^{1,*,†} , and Mario Sicuranza ^{1,†} 

¹ Institute for High Performance Computing and Networking-National Research Council of Italy (ICAR-CNR), 80131 Napoli, Italy; nadia.brancati@icar.cnr.it (N.B.); mario.sicuranza@icar.cnr.it (M.S.)

² Department of Diagnostic Imaging, University of Naples “Federico II”, 80138 Napoli, Italy; simona.depietro@unina.it

³ Pineta Grande Hospital, 81030 Castel Volturno, Italy

* Correspondence: teresa.conte@icar.cnr.it (T.C.); martina.russo@icar.cnr.it (M.R.)

† These authors contributed equally to this work.

Abstract

The widespread fragmentation of patient information across heterogeneous systems and the lack of standardized integration mechanisms hinder efficient and comprehensive medical diagnostics. To address these limitations, this work presents an architecture framework designed to support physicians in the diagnostic process by integrating clinical and socio-health information (patient medical histories), structured documents extracted from Health Information System (HIS), and data automatically extracted from diagnostic images using Artificial Intelligence (AI) techniques. The proposed architecture is made by several modules, in particular a Decision Support System (DSS) that enables risk assessment related to specific patient’s clinical conditions. In addition, the clinical information retrieved is aggregated, standardized, and transmitted to external systems for follow up. Standardization and data interoperability are ensured through the adoption of the international HL7 Fast Healthcare Interoperability Resources (FHIR) standard, which facilitates seamless connection with HIS. An Android application has been developed to communicate with different HISs in order to: (i) retrieve information, (ii) aggregate clinical data, (iii) calculate patient risk scores using AI algorithms, (iv) display results to healthcare professionals, and (v) generate and share relevant clinical information with external systems in a standardized format. To demonstrate architecture’s applicability, a case study on breast cancer diagnosis is presented. In this context, an AI-based Risk Assessment module was developed using the Breast Ultrasound Images Dataset (BUSI), which includes benign, malignant, and normal cases. Machine Learning algorithms were applied to perform the classification task. Model performance was evaluated using a 4-fold cross-validation strategy to ensure robustness and generalizability. The best results were achieved using the Multilayer Perceptron method, with a competitive F1-score of 0.97.

Keywords: FHIR; breast cancer; machine learning; ultrasound images; Decision Support System



Academic Editor: Phillip Olla

Received: 7 November 2025

Revised: 26 November 2025

Accepted: 30 November 2025

Published: 2 December 2025

Citation: Brancati, N.; Conte, T.; De Pietro, S.; Russo, M.; Sicuranza, M. Machine Learning-Enhanced Architecture Model for Integrated and FHIR-Based Health Data. *Information* **2025**, *16*, 1054. <https://doi.org/10.3390/info16121054>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, digital healthcare has witnessed an exponential growth in the volume and variety of clinical data generated from heterogeneous sources such as Health Information Systems (HISs), wearable devices, imaging systems, laboratories, and mobile applications. However, this abundance of information has not translated into effective

clinical integration: data often remain fragmented, inaccessible, or non-interoperable across different solutions, platforms, and companies [1–3].

To address these issues, HL7 Fast Healthcare Interoperability Resources (FHIR) (<https://www.hl7.org/fhir/>, accessed on 26 July 2025) has emerged as a modern and modular standard designed to support interoperable health data exchange through well-known web technologies such as RESTful APIs, JSON, and XML. FHIR organizes clinical information into granular and reusable “resources” (e.g., Patient, Observation, Condition), enabling flexible integration among HISs, clinical systems, and computational models [4], including those based on Artificial Intelligence (AI). Before the adoption of such standards, the absence of structured and semantically consistent data hindered the full potential of AI-based tools and Decision Support Systems (DSS), which rely on standardized inputs to produce reliable, reproducible, and personalized recommendations [5]. By standardizing diagnostic metadata and clinical information, FHIR facilitates the development of interoperable, reusable, and scalable DSS, accelerating knowledge transfer across diverse clinical environments [6]. According to a recent systematic review, over 98% of DSS tools developed between 2018 and 2021 adopted FHIR as the main interoperability standard [7].

AI-based DSS and Computer-Aided Diagnosis (CAD) tools can be used to support clinicians in their daily diagnosis process, for example, by automatically analyzing medical images. In the context of breast cancer, systems like the Automated Breast Ultrasound System (ABUS) have demonstrated advantages over handheld Ultrasound (US) images, enhancing lesion visualization and reducing operator dependency [8]. However, diagnosing breast cancer through US imaging remains a challenging task. Although the US is a non-invasive and widely available technique, its interpretation strongly depends on the operator’s expertise and is often influenced by inter-observer variability and image noise. Furthermore, the morphological similarity between benign and malignant lesions can make accurate discrimination difficult, particularly in dense breast tissues [9–11]. To address these challenges, Machine Learning (ML)-based CAD systems are being increasingly developed to support clinicians in the automatic analysis of US images [12–15]. Many studies use ML approaches, combined with Deep Learning (DL) models, aiming to obtain interpretable models, improving performance at the same time [16–19]. Other methods are based only on deep networks [20,21]. Existing FHIR-based architectures provide significant but partial contributions: some focus on interoperability and integration with EHRs, others illustrate how FHIR can enable AI-driven clinical workflows, while others define oncology-specific profiles.

Our research aims to improve the integration of AI-based DSSs into clinical practice by defining an architecture that incorporates several modules capable of communicating with each other and with heterogeneous external eHealth systems using the interoperable HL7 FHIR standard.

In this work, we demonstrate the use of the architecture through a case study related to the automated diagnosis of breast cancer via US image analysis based on AI approaches. Finally, we describe the patient-centered mobile application developed for the architecture, called InferCare.

The main innovations and distinctive aspects of our approach are:

- **Data Interoperability:** Existing clinical systems do not provide an integrated and patient-centered ecosystem that unifies anamnesis, structured clinical documents, diagnostic imaging, and AI-based decision support within a single workflow. The proposed architecture addresses this gap by integrating patient historical records, diagnostic imaging data, and risk assessment information originating from heterogeneous data sources. Interoperability is ensured through compliance with the HL7 FHIR standard, facilitating standardized data exchange and semantic consistency across HISs.

- Standard-driven AI integration: AI modules are integrated in compliance with recognized interoperability and transparency standards, ensuring that all AI-based results remain traceable, interpretable, and consistent with the overall system workflow.
- Bidirectional data flow: The clinical information collected from patients, encompassing medical history and diagnostic results, is semantically structured and encoded as HL7 FHIR resources. This approach enables standardized representation and supports bidirectional interoperability with heterogeneous external HIS, allowing both the retrieval of existing clinical data and the transmission of processed or newly generated information.
- User interaction: The InferCare application offers an intuitive interface that enables seamless interaction between patients and clinicians, effectively bridging mobile health (mHealth) solutions and institutional HIS.
- Clinical focus and validation: The interactions among the modules within the proposed architecture are demonstrated through a breast cancer case study based on US images analysis. This evaluation highlights the feasibility of integrating modules implementing ML algorithms with interoperability modules, thereby validating the system's capacity to support clinically relevant diagnostic workflows.

This combination of (i) interoperability, (ii) AI-based results, and (iii) patient-centered design represents a comprehensive and scalable architecture capable of bridging the gap between fragmented healthcare data and DSS.

2. Related Works

The urgent need to adopt standardized data models, such as the HL7 FHIR standard, has stimulated a wide range of initiatives addressing different aspects of healthcare data management, from interoperability with legacy systems to AI-driven clinical applications and oncology-specific profiling. However, these contributions are often fragmented, each addressing a specific challenge, such as data exchange, integration of mobile health applications, or definition of oncology data elements, without providing a comprehensive framework that unifies anamnesis, structured data (e.g., diagnostic imaging, medical reports, etc.), and DSS.

The use of the FHIR standard into HISs has favored the development of solutions aimed at professionals, enhancing data accessibility, quality, and interoperability.

Within the field of interoperability, one of the most influential initiatives is “SMART on FHIR” [22]. This approach proposes a modular framework that not only fosters interoperability but also facilitates the secure integration of third-party applications into HISs, and in particular into Electronic Health Records (EHRs). Through semantically constrained FHIR profiles, OAuth2 (<https://oauth.net/>, accessed on 12 September 2025) authorization, and OpenID Connect (<https://openid.net/>, accessed on 12 September 2025) authentication, the platform SMART on FHIR enables the development of reusable clinical applications, easily integrated into healthcare professionals' workflows.

However, despite its flexibility, the framework focuses mainly on application interoperability, without addressing semantic consistency or integration of heterogeneous data sources such as imaging and patient-generated content.

Another proposal that uses the FHIR standard in an integrated way within a framework architecture is Drishti [23] that extends the “Open mHealth” framework with a modular sense–plan–act architecture, designed to enable personalised behavioural interventions in mHealth. It connects data collection, planning, and alert delivery modules via RESTful APIs and FHIR-compatible backends by integrating FHIR resources such as Observation and CarePlan. This supports seamless integration with clinical systems such as “OpenMRS” (<https://openmrs.org/>, accessed on 12 September 2025).

Open mHealth uses FHIR as the canonical format for data exchange and storage, promoting interoperability, reusability, and modular development across different mobile health applications.

Nevertheless, this architectural proposal is primarily oriented toward data collection and monitoring patient behavior, offering limited support for complex clinical workflows and multimodal data integration.

The “MDIRA” initiative [24] offers a vendor-neutral, standards-based reference architecture for clinical device interoperability. Using IEEE 11073 semantics, IHE profiles, and HL7 FHIR messaging, MDIRA enables the seamless and secure integration of medical devices in hospital, home, and hospital-in-home settings. MDIRA supports both peer-to-peer and ICE-style (<https://mdpnp.mgh.harvard.edu/astra-portfolio/ice-standard-integrated-clinical-environment>, accessed on 26 July 2025) peer-to-aggregator communications and facilitates the development of autonomous, reliable, and reusable systems for critical care delivery.

Although MDIRA is effective in integrating medical devices, its scope remains strictly limited to devices in hospital, home, and hospital-in-home settings. It does not take into account other sources of clinical data, such as patient-generated information or diagnostic images, nor does it fully support complex clinical workflows.

Complementing these approaches, an ECG stream analysis framework [25] demonstrates how FHIR can support AI-driven healthcare applications in a cloud native environment. Using the Google Cloud Healthcare API, it securely stores FHIR-encoded ECG data and processes it using tools such as Scikit-Learn and PyTorch (<https://pytorch.org/>), thereby bridging the gap between clinical data interoperability and advanced AI real-time analytics and personalized monitoring.

The ECG Stream Analysis Framework effectively demonstrates how FHIR can enable AI-based analysis of real-time physiological data in a cloud-native setting, but its scope is limited to unidimensional ECG signals. It does not address multimodal integration, semantic standardization, or comprehensive decision-support workflows, which are instead central to the architecture proposed in this work.

In the field of precision medicine, the mCODE initiative [26] defines standardized FHIR profiles for oncology data, promoting reuse in both clinical systems and research. This model has been extended to international contexts [27], showcasing its adaptability to different healthcare systems. Similarly, the OSIRIS project [28] proposes a FHIR-compatible framework to enhance the sharing and analysis of clinical and genomic data in oncology.

Despite their contribution to oncology data standardization, these initiatives focus on defining specific FHIR profiles rather than proposing integrated architectures that connect clinical, imaging, and AI-driven components.

Major et al. [29] demonstrate how a FHIR back-end, integrated with the “Epic” system, which is one of the most widely used EHR systems in the world in hospital setting [30], allows real-time retrieval and analysis of clinical notes, medications, and vital signs to support AI-based models, enhancing timely clinical decision-making. Similarly to our approach, this demonstrates how FHIR-structured data can support decision-making processes across diverse clinical scenarios.

Although the literature demonstrates significant progress in standardizing healthcare data and promoting interoperability, particularly through the adoption of HL7 FHIR, existing solutions remain fragmented, each addressing only a portion of the clinical data ecosystem. This work fills this gap by proposing an end-to-end, FHIR-native architecture that consolidates heterogeneous patient information, embeds a modular DSS for risk assessment, and supports standardized communication with external HISs. Such an archi-

texture is necessary because it provides a unified, FHIR-native workflow that ensures the standardized and integrated data required for reliable AI-based clinical decision support.

3. Integrated Patient Decision Support System—Architecture and Methodology

This section illustrates the definition of the architecture of the Integrated Patient Decision Support System, namely IPDSS, a solution designed to centralize, integrate, and standardize heterogeneous patient information in order to provide effective decision support to healthcare professionals and facilitate integration with existing HISs. Figure 1 presents a comprehensive view of the proposed modular architecture.

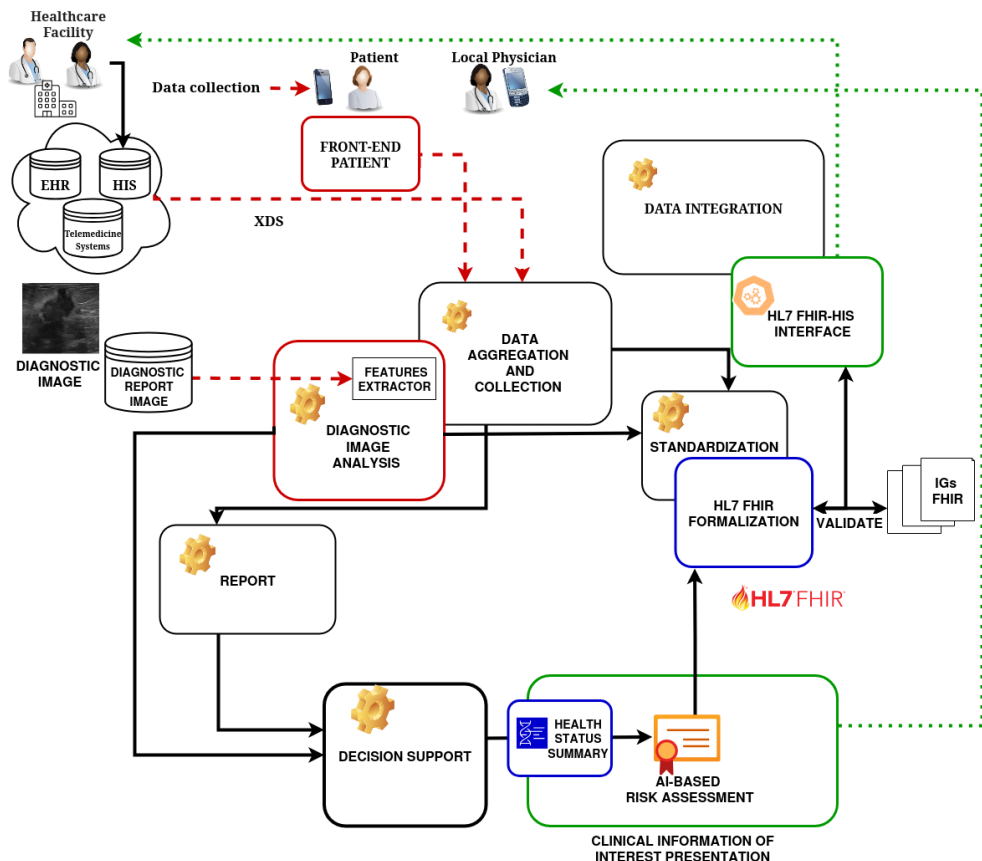


Figure 1. Architecture system. The front-end components are represented by colored boxes, while the overlapped black boxes indicate the back-end modules of the architecture. Input components are represented with red dashed arrows and red boxes, whereas output components use green dotted arrows and green boxes.

The architecture is designed as a modular, standards-based ecosystem that captures clinical and imaging data from multiple sources, applies algorithmic analysis and decision-support logic, and presents consolidated outputs via a user-centric interface. All information is formalized as HL7 FHIR resources and governed by a dedicated FHIR Implementation Guide, ensuring alignment with external HIS/EHR systems for seamless interoperability and future reuse.

The main objectives of IPDSS are:

- **Data Aggregation and Collection:** Aggregating information from different sources, such as (i) information collected from interviews or through forms filled in by patients (medical and family history); (ii) data derived from diagnostic image analysis systems and reports; (iii) structured data obtained from HIS.

- **Data Standardization:** Formalizing all data (source, integrated, and processed) using the international HL7 FHIR standard to realize an interoperable solution.
- **Intuitive and fast visualization of interest information:** Providing integrated information of interest for healthcare professionals through an easy-to-navigate and understand user interface. Only information deemed useful is displayed for the specific diagnostic case.
- **Decision Support:** Generating a concise summary of the patient's health status based on the integrated data and the use of AI algorithms to propose an AI-based risk assessment.
- **Data Integration and Communication:** Following data processing and integration, the architecture builds and manages HL7 FHIR-compliant resources that represent patient information, diagnostic results, and derived information. These resources are then sent to external systems such as HIS and EHR, contributing to the enrichment of patient clinical information and the overall increase in clinical knowledge related to patient health status.

In the following, we go into the details of the architecture's description, with reference to functional requirements, data model, data flow, and Diagnostic Image Analysis Module.

3.1. Architecture

The proposed solution adopts a modular architecture based on logically separated modules or components that communicate with each other.

In Figure 1, the front-end components are represented by colored boxes, while the overlapped black boxes indicate the back-end modules of the architecture. Input components are represented with red dashed arrows and red boxes, whereas output components use green dotted arrows and green boxes.

The front-end modules include:

- **Front-end Patient**
It allows the patient to independently collect a range of anamnestic information (possibly specific to the clinical condition). The module allows for the management of data obtained by a form presented to the patient for the collection of medical history and other relevant information (allergies, medication, family history, symptoms, etc.). This module also allows the patient to retrieve any information present in the system through interaction with the HL7 FHIR-HIS interface module.
- **Diagnostic Image Analysis Module (DIAM)**
It performs processing and analysis of diagnostic images (e.g., detection, segmentation, classification, feature extraction) using ML algorithms or specific analysis tools. The aim of this module is to return an AI-based risk assessment.
- **HL7 FHIR Formalization Module (HFFM)**
This module allows for formalizing, through international standard HL7 FHIR profiles, the data provided by the patient during the anamnesis phase (coming from the patient's anamnesis form), and the results of image analysis (from the DIAM). This module uses the FHIR profiles and resources (e.g., Patient, Observation, Diagnostic Report, Allergy Intolerance, Medication Statement), which are appropriately defined to ensure compliance with FHIR standards. It manages the creation of FHIR resources and carries out validation for the aim of proposed architecture.
- **Clinical Information of Interest Presentation Module (CIIPM)**
This module allows identifying and collecting all and only the clinical information of interest for a specific diagnosis. It thus enables the presentation of an integrated and intuitive view of the information: personal data, structured medical history, and image analysis results.

- **Health Status Summary Module**
It applies logical rules and algorithms to extract and analyze integrated data (history, image results, other available data), and generates a concise summary of the patient's health status, highlighting key information, potential risks (AI-based risk assessment), and recommendations.
- **HL7 FHIR-HIS Interface**
It enables the CIIPM to communicate with the existing platform (EHR) using the HL7 FHIR standard and supports FHIR operations such as Create, Read, Update of Patient resources, Observation, Diagnostic Report, and other relevant ones.

The workflow illustrated in Figure 1 begins with the collection of data from multiple sources and proceeds through aggregation, analysis, standardization, and the presentation of results, ensuring interoperability via HL7 FHIR. Patients actively contribute by providing personal and clinical information through the system, including structured anamnesis, demographic data, family medical history, symptoms, allergies, medications, lifestyle factors, consent documentation, and optional clinical attachments. Physicians access patient data either entered directly by patients or retrieved from HIS/EHR systems via an interface that communicates through XDS with external HIS/EHRs. Additional input data come from diagnostic images associated with the patient, processed by the DIAM using ML algorithms, and from structured documents available in HIS/EHR systems. These documents may be generated by physicians in external systems and retrieved through the HL7 FHIR interface (XDS). The decision support module generates a concise summary of the patient's health status based on integrated data and AI algorithms, providing an AI-driven risk assessment. Results are presented to physicians through the CIIPM, displaying only the most relevant clinical information, including integrated medical history, image results, and risk scores. All collected data are processed by the HFFM, which converts them into FHIR resources (e.g., patient, observation, diagnostic report) and validates them according to the relevant Implementation Guides (IGs). Finally, standardized FHIR data are transmitted to external HIS/EHR systems to support data enrichment and follow-up, completing the care cycle with healthcare facilities and local physicians.

Data Model

The data used in the architecture presented in the previous section are anamnestic data, i.e., a collection of information about the patients and their medical history, carried out by the physician to better understand the situation and make an accurate diagnosis. These data help the physician to identify possible diseases, risk factors, and family predispositions, which are essential for appropriate treatment. In addition to the medical history data, some data related to patient images are provided to the clinicians to support the diagnosis (Pathological Features). Moreover, other data features are also extracted from the images (Hand-crafted Features) that allow the computation of the AI-based risk assessment. Section 4 explains the process for the computation of AI-based risk assessment.

3.2. Architecture Modules Details

The following paragraph provides the flow of data within the proposed *IPDSS* architecture.

3.2.1. Front-End Patient

This module is designed to enable patients to autonomously provide a wide range of anamnestic information, which can be tailored to their specific clinical condition. Through an intuitive user interface, patients are presented with a customizable form where they can input relevant data regarding their personal and family medical history. This includes, but

is not limited to, information about existing or past symptoms, current medications, known allergies, past diagnoses, and hereditary conditions.

The data can be collected either directly through the dedicated mobile application or remotely during teleconsultation interviews with physician, ensuring flexibility and accessibility for different clinical contexts.

The collected data are systematically managed and stored by the module, ensuring that healthcare professionals can access accurate and up-to-date patient information. Furthermore, the module is integrated with HIS/EHR, allowing patients to retrieve any relevant data already present in the system.

This bidirectional communication ensures that both patient-provided information and existing clinical records are seamlessly synchronized, enhancing the completeness and reliability of the patient's health profile.

3.2.2. Diagnostic Image Analysis Module (DIAM)

DIAM aims to provide two types of support to the healthcare professionals:

- Some specific features related to the pathology, namely Pathological Features (PFs), are automatically extracted from the images by using properly designed Computer Vision algorithms. These features are shown to the healthcare professionals and are useful to support the diagnosis.
- A risk assessment, computed by using suitably designed ML algorithms. For the computation of AI-based risk assessment, the ML methods use a set of features automatically extracted from US images, larger than PF, and called Hand-crafted Features (HFs).

3.2.3. HL7 FHIR Formalization Module (HFFM)

HFFM transforms clinical and contextual data collected by the user into resources that comply with the HL7 FHIR standard. Specifically, the module processes three main categories of information:

- Anamnestic data: information provided directly by the patient through the Front-end Patient module. This includes, for example, self-reported symptom medical conditions, self-reported symptoms, ongoing treatments, and relevant lifestyle factors.
- Clinical data derived from structured documents: information provided by external systems such as HIS/EHR.
- Clinical information derived from diagnostic images: includes results obtained through the automated analysis of diagnostic images (e.g., US images) by the DIAM. This information is then translated into formalized clinical observations according to the HL7 FHIR standard.

In line with international approaches that enable information management through the FHIR standard and the formal definition of a project-specific Implementation Guide (e.g., [31]), which specifies the rules and constraints governing the interoperable use of health data and information, we adopted the IG presented in [32]. In the proposed IG, profiles, value sets, and system codes were defined and formalized. The developed IG also incorporates concepts from internationally recognized guides, such as the International Patient Summary (IPS) [33], which establishes an interoperable structure for summarizing a patient's clinical information, thereby facilitating the exchange of essential health data. Within the proposed architecture, the IG defined in [34,35] is leveraged to support the structuring of the information used by our solution. To guarantee a uniform mapping between the heterogeneous input data and the HL7 FHIR-based representation, a structured correspondence was defined. Table 1 summarizes this mapping, linking each clinical concept with its data source and the corresponding FHIR profile used in the IG. The information

originates from various contexts, such as IPS CDA 2.0 sections, patient interviews, or the DIAM, and is semantically aligned with the related FHIR resources.

Table 1. Mapping between concepts, source data, and FHIR profiles.

Concept to Be Mapped	Source Data	FHIR Profile
Allergy or Intolerance	IPS CDA 2.0	AllergyIntolerance_Patient
Appointment	Physician interview and source app	Appointment_Patient
Risk Assessment	DIAM	CancerRiskAssessment_Patient
Care Plan	IPS CDA 2.0	CarePlan_Patient
Condition	IPS CDA 2.0	Condition_Patient
Consent of Patient	App and EHR source	Consent_Patient
Diagnostic Report	IPS CDA 2.0	DiagnosticReport_Patient
Family Member History	IPS CDA 2.0 and physician interview	FamilyMemberHistory_Patient
Medication	IPS CDA 2.0	Medication_Patient/ MedicationStatement_Patient
Feature Major Axis, Minor Axis	DIAM	Observation_Axis
Observation (symptoms, etc.)	IPS CDA 2.0	Observation_Patient
Feature Orientation	DIAM	Observation_Orientation
Feature Circularity	DIAM	Observation_Circularity
Feature Perimeter Regularity	DIAM	Observation_PerimeterRegularity
Patient	Patient app	Anamnesis_Patient
Physician	App or interview	Practitioner_Anamnesis
Procedure	IPS CDA 2.0	Procedure_Patient

In addition, to exemplify the transformation process carried out by the HFFM, Table 2 illustrates the detailed correspondence between specific CDA 2.0 elements and the equivalent FHIR elements for the AllergyIntolerance resource. This mapping provides a concrete example of how information from CDA-based documents is converted into structured FHIR resources by the formalization module.

3.2.4. Clinical Information of Interest Presentation Module (CIIPM)

This module is responsible for identifying and aggregating all the clinical information that is specifically relevant to a given diagnosis, based on the patient's current clinical condition and medical history. Its goal is to filter out non-essential data and focus only on what is diagnostically significant, thereby reducing information overload for health-care professionals.

The module presents this curated information through an integrated and user-friendly interface. It combines various data sources, including personal and demographic information, structured and categorized medical history, and the results obtained from medical image analysis, into a coherent and accessible overview. This facilitates faster and more informed clinical decision-making by providing a comprehensive yet focused snapshot of the patient's health status.

Table 2. Mapping between CDA elements and FHIR AllergyIntolerance resource attributes.

CDA	FHIR
ClinicalDocument/structuredBody/component/section/id	AllergyIntolerance/identifier
ClinicalDocument/structuredBody/component/section/text	AllergyIntolerance/text
ClinicalDocument/structuredBody/component/section/entry/act/statusCode	AllergyIntolerance/status
ClinicalDocument/structuredBody/component/section/entry/act/effectiveTime	AllergyIntolerance/recordedDate
ClinicalDocument/structuredBody/component/section/entry/act/entryRelationship/observation/code	AllergyIntolerance/type
ClinicalDocument/structuredBody/component/section/entry/act/entryRelationship/observation/effectiveTime	AllergyIntolerance/onset
ClinicalDocument/structuredBody/component/section/entry/act/entryRelationship/observation/participant/participantRole/playingEntity/code	AllergyIntolerance/substance
ClinicalDocument/structuredBody/component/section/entry/act/entryRelationship/observation/entryRelationship/observation/effectiveTime	AllergyIntolerance/reaction/onset
ClinicalDocument/structuredBody/component/section/entry/act/entryRelationship/observation/entryRelationship/observation/value	AllergyIntolerance/reaction/manifestation
ClinicalDocument/structuredBody/component/section/entry/act/entryRelationship/observation/entryRelationship/observation/value	AllergyIntolerance/reaction/certainty
ClinicalDocument/structuredBody/component/section/entry/act/entryRelationship/observation/entryRelationship/act/text	AllergyIntolerance/note

3.2.5. Health Status Summary Module

This module performs a synthesis of the patient’s overall health condition by applying predefined logical rules and advanced algorithms to the available integrated data (diagnostic image analysis results, and any other clinically relevant data available in the system).

Based on this analysis, the module generates a clear and concise summary of the patient’s health status. It highlights key clinical findings, flags potential health risks through AI-based risk assessment models, and provides actionable recommendations when appropriate. This summary is designed to support healthcare providers by offering a quick yet comprehensive overview, aiding in both diagnosis and treatment planning.

3.2.6. HL7 FHIR—HIS Interface

This module facilitates seamless communication between the CIIPM and the existing HIS, such as the EHR platform. It leverages the HL7 FHIR standard, which is widely adopted for the secure and efficient exchange of healthcare data.

The interface supports a range of core FHIR operations, including the creation, retrieval, and updating of key healthcare resources such as Patient, Observation, DiagnosticReport, and other relevant resource types. By implementing these operations, the module ensures interoperability with other systems and enables real-time synchronization and

sharing of clinical information across the healthcare infrastructure. This interoperability enhances the consistency, accuracy, and accessibility of patient data throughout the clinical workflow.

4. Case Study: Breast Cancer

This work proposes a case study related to the clinical condition of breast cancer.

In the following section, we describe the modules, DIAM and HFFM, that require customization in relation to the selected case study. The use of the specific scenario allowed us to validate the definition of the entire architecture, showing the relationships between the various components and therefore the flow of heterogeneous data, its use and processing in order to obtain a result (risk assessment) in terms of summarized information produced and how this information can be standardized and sent to external systems using the FHIR standard.

4.1. Implementation of DIAM in the Breast Cancer Case

- Dataset description

In this work, the Breast Ultrasound Images Dataset (<https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset>, accessed on 19 March 2025) (BUSI) is used for the evaluation of the classification of US images. The dataset includes 780 breast US images of women aged between 25 and 75 years old. A total of 210 images contain malignant lesions, 437 are annotated as benign breast cancer, and 133 are normal breast cancer images. Only malignant and benign cases are included in this study. In detail, for each US image, a ground truth with the annotated lesion is associated.

As an example, Figure 2 illustrates two breast US images: one showing a malignant lesion and the other a benign one. In both cases, the lesions are contoured in red by the expert radiologists, which allowed us to extract the key morphological and color-related features directly in that annotated portion and use them for the classification process.

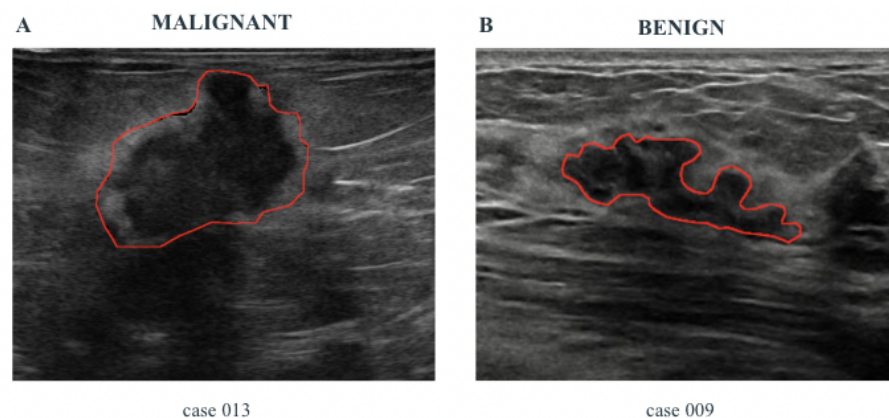


Figure 2. US breast images. Example of two US images of the breast, one with a malignant lesion (A) and one with a benign lesion (B). The images show the lesions contoured in red by the expert radiologists.

- AI-based Risk assessment calculation

For the selected case study, the DIAM provides the PF set, reported in the Table 3. These features are among the most used by radiologists to provide a diagnosis of breast cancer [36]. So, we show them to the clinician to support the diagnostic decision process.

Table 3. Description of PF set.

Feature	Description
Major Axis, Minor Axis	Provides intuitive size and shape information
Perimeter Regularity	Irregular contours may be indicative of malignancy
Orientation	An angle greater than 45° may suggest a malignant nature
Circularity	Lower circularity can reflect irregular lesion shapes

For the AI-based risk assessment, a classification of the lesion as benign or malignant is proposed. For the classification purpose, we chose to use handcrafted features. Unlike DL approaches, the use of handcrafted features has the advantage of producing an interpretable model that can show the clinician which features were used to achieve that result. Thus, we use the HF set, composed of morphological and color-related features, reported in Table 4, extracted from the annotated lesions.

Table 4. Description of HF set.

Feature	Description
Eccentricity	Indicates how elliptical the lesion is (0 = perfect circle, 1 = highly elongated ellipse)
Circularity	Measures how similar the lesion is to a circle (1 = perfect, <1 = more irregular)
Perimeter Regularity	Evaluates the complexity of the lesion boundary
Axis Ratio	Relationship between the axes of an ellipse or ellipsoid
Solidity	Ratio between the actual area and the convex area (1 = compact, <1 = irregular)
Extent	Ratio between the lesion area and its bounding box
Elongation	Indicates whether the lesion is stretched along a particular direction
Fractal Dimension	Quantifies the complexity of the lesion's contour
Area	Number of pixels comprising the lesion
Perimeter	Length of the lesion's contour
Convex Area	Area of the convex hull surrounding the lesion
Equivalent Diameter	Diameter of a circle having the same area as the lesion
Kurtosis	Measures the “peakedness” of the intensity distribution
Skewness	Measures the asymmetry of the intensity distribution
Entropy	Indicates the degree of randomness in the pixel intensity distribution
Contrast	Measures local intensity variation
Homogeneity	Quantifies how similar neighboring pixels are

A feature selection step is applied in order to assess the HFs that are more important for the classification. Recursive Feature Elimination (RFE) using Logistic Regression (LR) results in being the best method for the feature selection phase [37]. Subsequent to the feature selection phase, a classification pipeline is performed with the objective of assessing the predictive capabilities of various ML algorithms. Specifically, four well-established classifiers are considered: Decision Tree (DT) [38], Multi-Layer Perceptron (MLP) [39], Naive Bayes (NB) [40], Random Forest (RF) [41]. Model performance is evaluated using

a four-fold cross-validation approach, where the dataset is randomly divided into four equally sized folds. At each iteration, three folds are used for training and one for testing.

In order to mitigate the issue of class imbalance inherent in the dataset, the Synthetic Minority Oversampling Technique (SMOTE) [42] is applied. Unlike traditional random oversampling methods, which tend to increase the risk of overfitting by merely duplicating existing minority class instances, SMOTE addresses the imbalance by synthetically generating new samples. This is achieved through interpolation between existing minority class samples that are in close proximity within the feature space. Such an approach fosters a more balanced class distribution and enhances the generalization ability of the trained models. For each image, the predicted class is the one associated with the highest probability score assigned by the ML algorithm. The risk assessment proposed to the clinician consists of the predicted class (benign or malignant) and the associated probability for that class.

Typically, clinicians interpret risk scores by mapping the numerical output of the AI model to well-defined clinical categories that reflect the likelihood of a specific condition, thereby guiding subsequent decision-making steps. In detail, clinicians combine three complementary elements: (i) the predicted class (benign or malignant) and its associated probability; (ii) the set of handcrafted features that contributed to the prediction, such as perimeter regularity, axis ratio, solidity, and circularity, which correspond to well-established radiological criteria; and (iii) the clinical context derived from the patient anamnesis and structured data. This interpretability enables clinicians to validate the AI output against their own visual assessment.

- Results

In Table 5, the obtained results, together with the standard deviation across the folds, are reported. They highlight the importance of carefully selecting discriminative handcrafted features for breast lesion classification from US images. The use of RFE proved particularly effective in identifying the most informative subset of features, suggesting that not all morphological descriptors contribute equally to the discrimination task. The fact that only three features—perimeter regularity, axis ratio, and solidity—were sufficient to achieve high performance indicates that these characteristics capture complementary aspects of lesion morphology that are highly relevant for distinguishing benign from malignant patterns. This is also consistent with radiological practice, where border irregularity, asymmetry, and spiculation are well-established hallmarks of malignancy.

Table 5. Results of the different ML algorithms used with the three key features. In bold, the best results for each measure.

	Accuracy	Precision	Recall	F1-Score
DT	0.9440 ± 0.028	0.9333 ± 0.027	0.9439 ± 0.028	0.9377 ± 0.028
MLP	0.9767 ± 0.008	0.9746 ± 0.008	0.9729 ± 0.008	0.9736 ± 0.008
NB	0.9488 ± 0.003	0.9596 ± 0.003	0.9254 ± 0.003	0.9397 ± 0.003
RF	0.9720 ± 0.015	0.9690 ± 0.014	0.9683 ± 0.015	0.9684 ± 0.015

Among the classifiers, MLP outperformed the others, achieving almost perfect discrimination with an accuracy close to 98% and a F1-score of 97%. This reinforces the idea that neural networks, even in relatively simple architectures, are well suited to model nonlinear feature interactions. While perimeter regularity or solidity alone can already provide meaningful information, their joint contribution, along with axis ratio, can be more effectively exploited by MLP compared to other models. In contrast, NB, constrained by its

independence assumption, showed limitations in handling such interactions, which likely explains its relatively lower recall.

DT performed adequately but was more affected by the dataset size and potential noise. This behavior is expected, as DTs are prone to overfitting when trained on small datasets and may fail to generalize well. RF, on the other hand, mitigated some of these issues by averaging multiple DTs, which resulted in solid performance (97.2% accuracy and 96.8% F1-score). However, RF still did not surpass MLP, suggesting that ensembles of shallow learners might not capture subtle, higher-order nonlinear relationships as effectively as neural networks.

Overall, these findings suggest that the careful combination of feature selection and the use of flexible learning models such as MLP can yield highly accurate breast cancer classification systems, even when relying solely on handcrafted features rather than deep representations. This is a particularly relevant result for settings where computational resources are limited or where large annotated datasets required for DL are not available. At the same time, the high performance obtained with a small feature set also improves interpretability, since the decision-making process can be directly related to clinically meaningful morphological descriptors.

4.2. Implementation of HFFM in the Breast Cancer Case

This section presents FHIR profiles developed to ensure the structured and interoperable representation of clinical and anamnestic information collected through an IG. Derived from the adaptation of international FHIR resources to the project context, these profiles cover several areas ranging from patient history to oncology risk analysis. In Table 6, a summary of a single profile of IG is reported, while in Figure 3, an overview of the Information Model and Profiled FHIR Resources is presented.

Table 6. Description of profiles.

Profile	Description
AllergyIntolerance_Patient	Used to represent the patient's known allergies and intolerances
Anamnesis_Patient	Used to represent the personal identity and demographic information of the patient subject to the anamnesis
Appointment_Patient	Used to describe information about scheduled clinical appointments
CancerRiskAssessment_Patient	Used to represent the patient's risk of developing cancer based on anamnesis and extracted clinical features
CarePlan_Patient	Used to represent a patient care or treatment plan
Condition_Patient	Used to represent a particular clinical condition of the patient
Consent_Patient	Used to represent the patient's informed consent regarding the use and sharing of their clinical data
DiagnosticReport_Patient	Used to document diagnostic reports associated with the patient, such as US images
FamilyMemberHistory_Patient	Used to document the patient's family medical history, with special reference to inherited diseases in family members
Medication_Patient	Used to describe the characteristics of drugs such as active ingredient, pharmaceutical form, and dosage
MedicationStatement_Patient	Used to represent the set of medications taken by the patient
Observation_Axis	Used to represent the orientation axis of a clinical image or structure, extracted from US or imaging data.

Table 6. Cont.

Profile	Description
Observation_PerimeterRegularity	Used to represent the regularity of the perimeter of a lesion or structure observed in diagnostic imaging
Observation_Orientation	Used to represent the spatial orientation of a lesion as detected in clinical imaging.
Observation_Circularity	Used to represent the circularity of a lesion or structure, derived from diagnostic image analysis
Observation_Patient	Used to record observations of the patient's health status such as vital parameters, symptoms, etc.
Practitioner_Anamnesis	Used to represent the health professional involved in patient care, e.g., the physician in charge of the examination.
Procedure_Patient	Used to represent medical procedures undergone by the patient, such as surgeries, biopsies, or diagnostic interventions

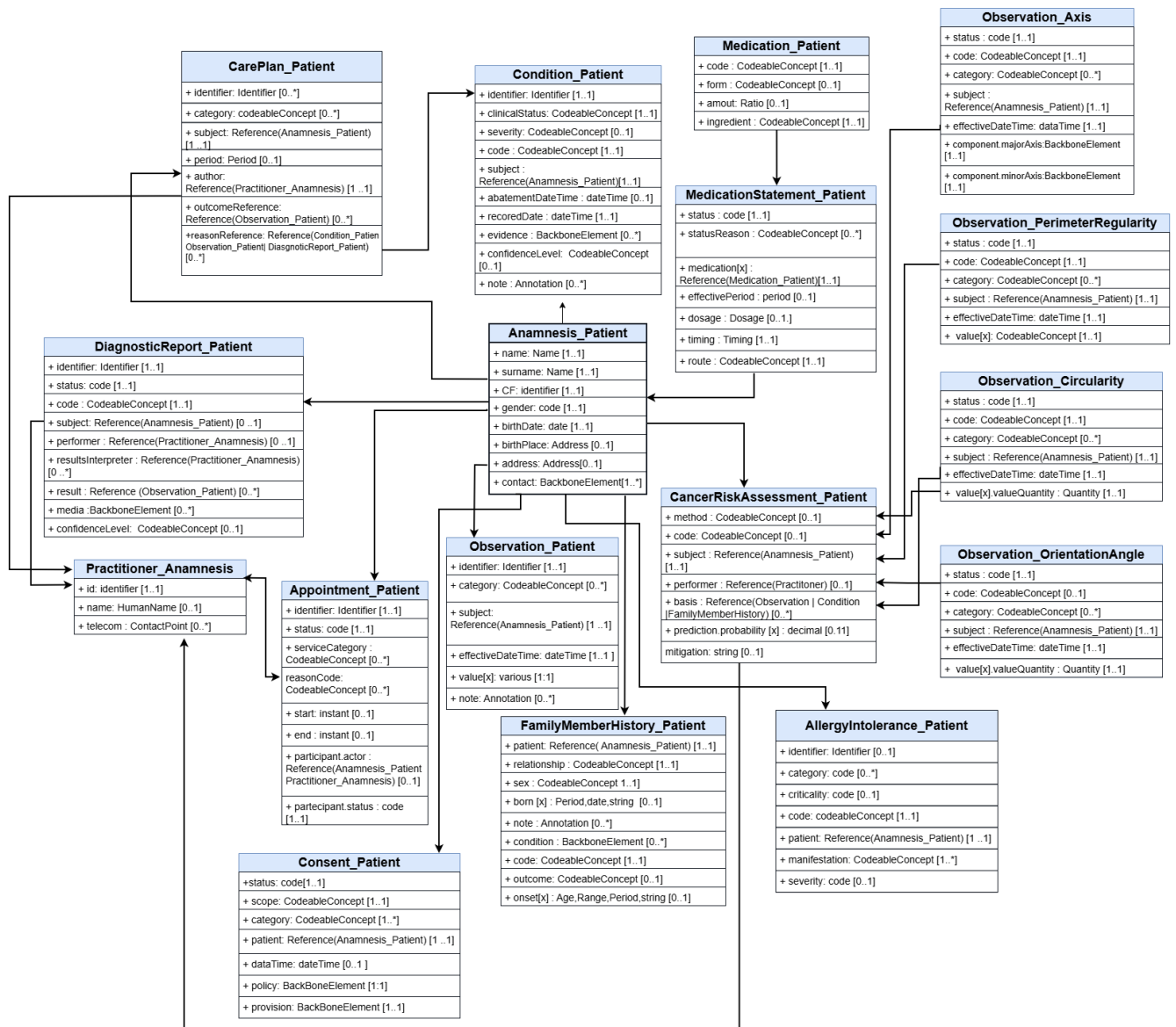


Figure 3. Overview of the Information Model and Profiled FHIR Resources. The asterisk “*” indicates that the element can have multiple occurrences.

5. InferCare Android Application

To facilitate interaction with the IPDSS system by both patients and physicians, the InferCare mobile app was developed. The app's name is derived from the combination of "Inference" (referring to AI inference) and "Care" (patient care). This section explains the interface and main features of the app, which allow guided completion of the medical history form, display of structured information, and support of diagnosis through summary views for physicians.

5.1. IPDSS Functional Requirements

The section specifies the Functional Requirements (FRs) of the IPDSS, organized according to a progressive nomenclature (FR01, FR02, etc.). These requirements, as you can see in the Table 7 below, define the expected behaviors of the system, both in terms of secure data acquisition and management and in terms of accessibility, user interface, and interoperability with external systems (such as EHR, HIS, and telemedicine systems).

Table 7. Functional requirements of the IPDSS.

Requirement ID	Description
FR01	IPDSS should enable patients to complete a structured form capturing personal data, teleconsultation history, family medical history, allergies, current medications, lifestyle habits, ongoing symptoms, and other clinically relevant information necessary for initial assessment.
FR02	The form should support multiple input types, including free text fields, multiple-choice options, checkboxes, and date pickers, to ensure flexibility and completeness of data collection.
FR03	IPDSS should implement client-side validation mechanisms to ensure data consistency, accuracy, and completeness before submission.
FR04	IPDS should allow patients to save their progress during form completion and resume the process at a later time without data loss.
FR05	IPDSS should ensure the confidentiality and integrity of the submitted information during data transmission through secure communication protocols (e.g., HTTPS, encryption).
FR06	A secure user authentication mechanism should be provided (where applicable) to control access to the form, particularly when enabling partial form saving or editing features.
FR07	IPDSS should be responsive and accessible across various devices, including tablets and smartphones, ensuring usability and inclusivity.

5.2. InferCare

The mobile app, InferCare, is designed to provide an intuitive, interactive interface that supports patients and physicians throughout the entire information flow. From the patient's perspective, an interface is implemented that enables the guided completion of a multi-section medical history form, covering symptoms, medications, allergies, and habits. The form can be temporarily saved locally and can be resumed later. Once the information is validated through the HFFM module, it is visible to the patient through the interface.

From the physician's perspective, an interface is implemented that provides authenticated and secure access to the patient dashboard, serving as a centralized entry point for all clinical information. Once logged in, the physician can view the list of waiting patients, available appointments, and a summary of each patient's health status. The Figures 4 and 5 show both the patient's and physician's viewpoints from the app interface. In the next section, the Sequence Diagrams for patient and physician interaction with the app are described.

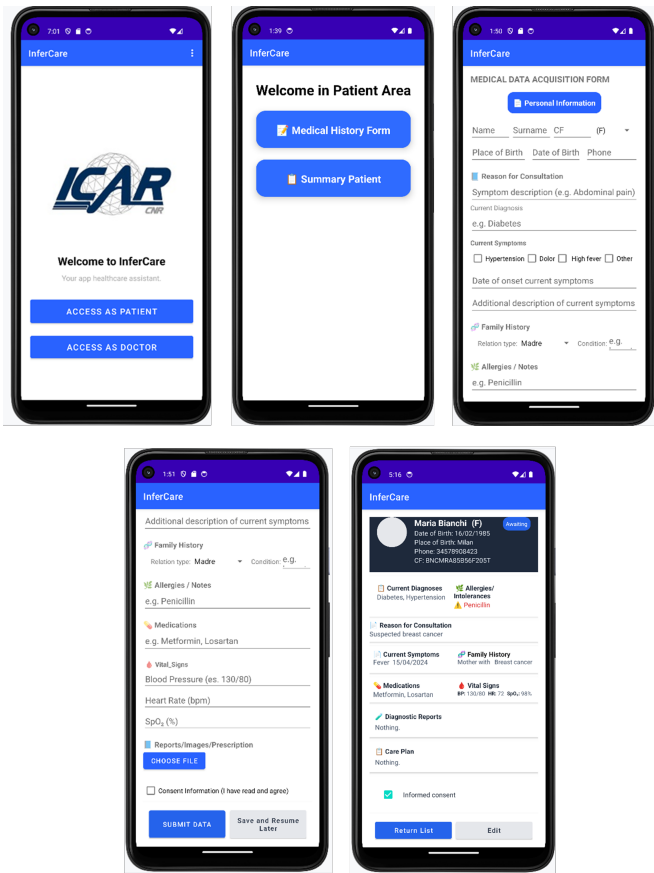


Figure 4. Patient’s view.

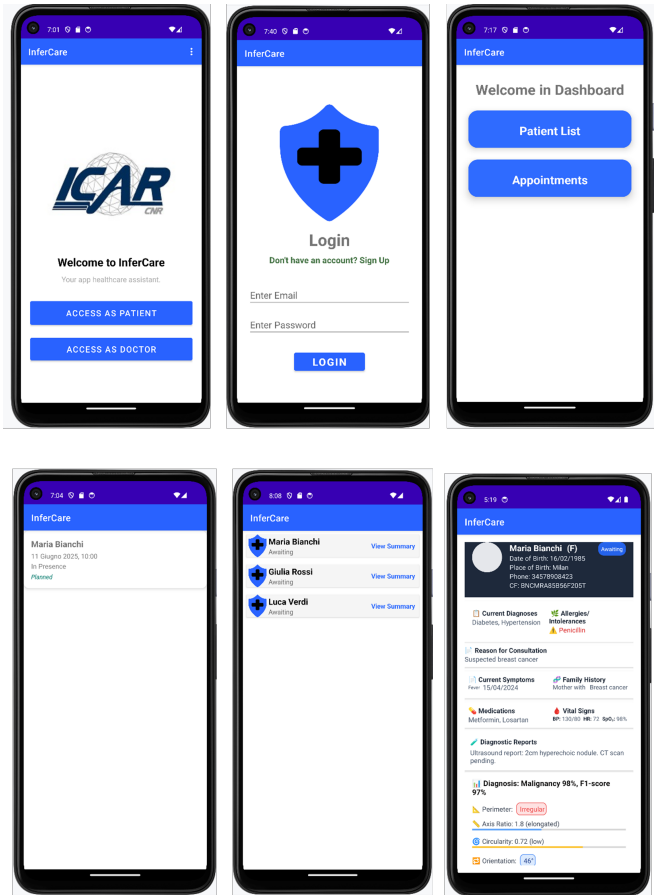


Figure 5. Physician’s view.

5.3. User Actions and Interactions

The sequences of actions performed by the patient are shown in Figure 6:

1. The patient starts the mobile application (StartApp);
2. A request is sent to the Front-end Patient to access the data entry form (RequestCompilationForm);
3. The patient completes a digital form through the Front-end Patient interface;
4. The Front-end Patient form sends the data to the Backend via a REST API (POST(data anamnesis module));
5. The “RESTful API” forwards the “POST (data anamnesis module)” to the “HL7 FHIR Formalization Module”;
6. The HL7 FHIR formalization module transforms the received data into FHIR resources and sends them to the “FHIR Server with IG” for data validation (“validate data”);
7. FHIR Server with IG validates and stores the FHIR Resources, then data send a “Response FHIR Resources” back to the HL7 FHIR Formalization module;
8. Finally, the “Front-End Patient” receives a “Response Resource FHIR” back to the mobile application as View Form anamnesis.

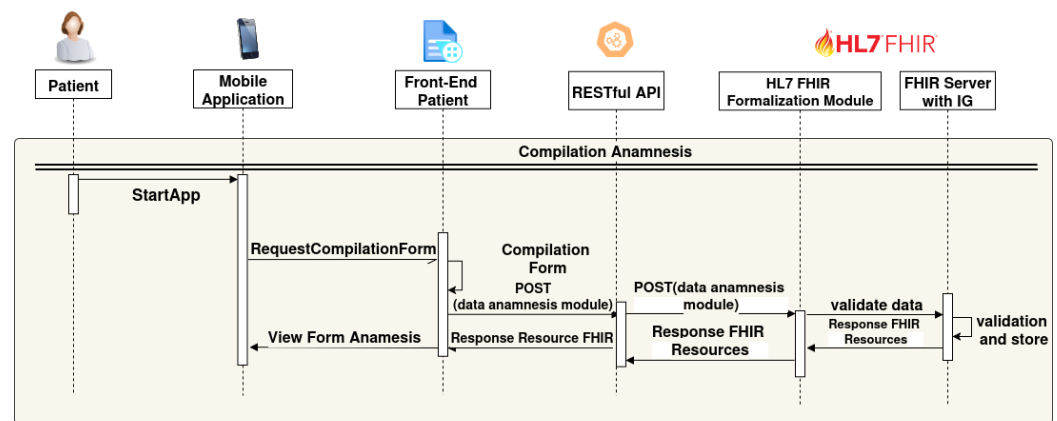


Figure 6. Sequence diagram, patient's view.

The sequence in Figure 7 describes the back-end process triggered when a user accesses the mobile application and requests the analysis of diagnostic images. The system handles image retrieval, clinical feature extraction via AI, and data transformation into HL7 FHIR format using a structured pipeline supported by RESTful communication and standardized data representation.

1. The user accesses the mobile application (AccessApp).
2. The application sends an image analysis request to the backend via RESTful API (RequestElaborationImage).
3. The RESTful API performs a POST request to the module that queries the US image database.
4. The database receives the request (RequestDiagnosticImage), retrieves the required US images (Images recovery), and returns them.
5. The US images are sent to the AI-based image analysis module.
6. The AI module extracts the clinical features from the ultrasound images (extract features images).
7. The extracted data are sent to the HL7 FHIR formalization module (send data features).
8. The HL7 FHIR module validates and transforms the data into FHIR format (validate data).
9. Finally, the FHIR Server with IG stores the validated features (validation and store features images resource).

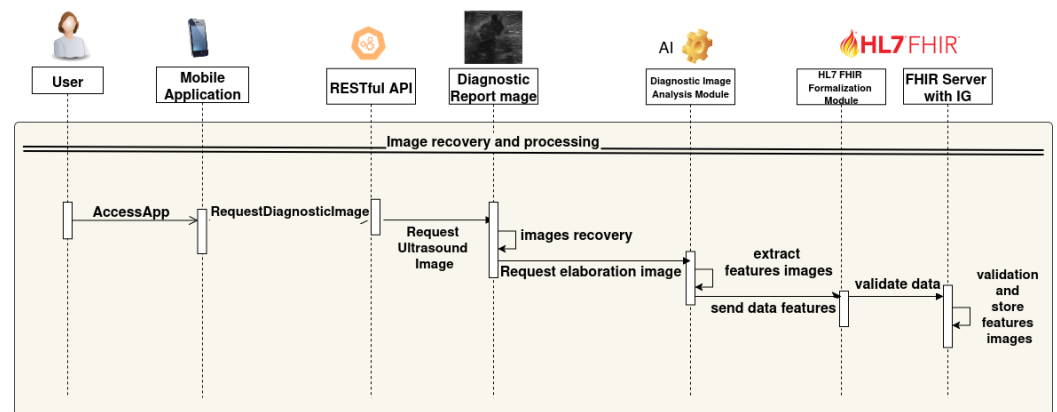


Figure 7. Sequence diagram image analysis.

The sequences of actions performed by the physician are as follows and shown in Figure 8:

1. The physician interacts with the mobile app to request the summary view (RequestViewSummary).
2. The app forwards the request to the Clinical Information of Interest presentation module (RequestSummary).
3. This module queries the FHIR Server through the RESTful API with a query containing the relevant data (Query FHIR Resources (DataAnamnesis, featuresImage)).
4. The FHIR Server with IG retrieves the requested clinical information and returns the FHIR Resources.
5. The Clinical Information of Interest presentation module extracts the main data from the received resources, creates a summary with a Health Status Summary and AI-based risk assessment, and highlights the principal information;
6. The summary view is presented to the physician through the mobile interface (View Summary).

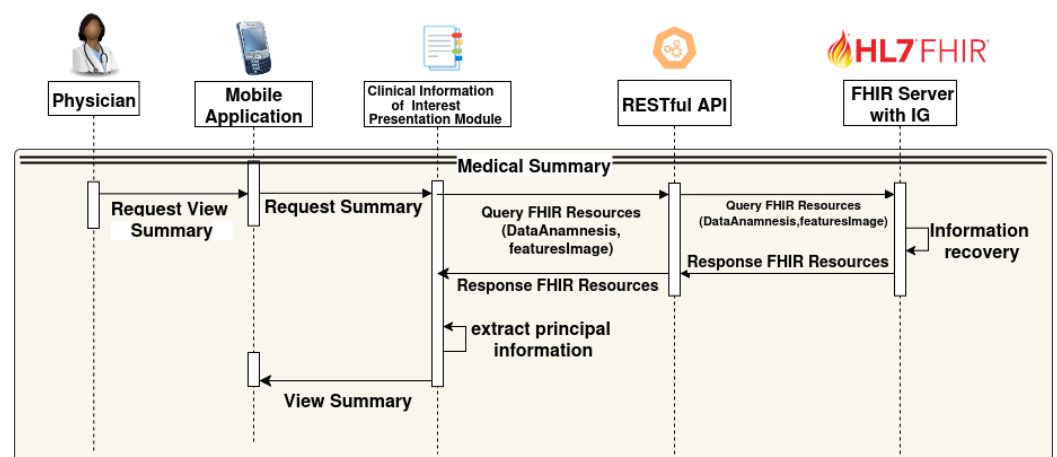


Figure 8. Sequence diagram, physician's view.

5.4. Data Flow Within the Architecture Enabled by the Developed InferCare

This section presents the flow of data within the proposed architecture, illustrated in Figure 9, showing how the *InferCare* mobile application interacts with heterogeneous HIS to ensure the interoperability, traceability, and semantic consistency of clinical data. The architecture integrates medical information from various sources, including EHRs, HISs, PACS, and patient-collected data, while adhering to international standards such as IHE-XDS, IHE-XDS-I, and HL7 FHIR. This multilevel integration enables structured access

to, and the secure exchange and validation of clinical and imaging information. Ultimately, it generates a unified and interoperable health summary that is accessible to physicians and patients. The operational workflow, detailed in Figure 9, unfolds as follows.

1. **Patient data collection:** the process begins with the patient completing a digital medical history form through the InferCare mobile application. The guided interface enables the structured collection of personal and clinical information, including demographic data, symptoms and reasons for consultation, family and medical history, allergies, medications in use, informed consent, and the upload of relevant clinical documents.
2. **Retrieval of Clinical Documents (EHR–IHE-XDS):** the InferCare application interfaces with the EHR through the IHE Cross-Enterprise Document Sharing (XDS) standard. Using the Registry Stored Query [ITI-18] IHE transaction, the app queries the XDS Registry to identify structured clinical documents associated with the patient. The registry provides the Query Response, allowing the system to retrieve the corresponding document set (Retrieve Document Set [ITI-43]) from the XDS Repository.
3. **Integration with Diagnostic Imaging Systems (IHE XDS-I/PACS):** when a new diagnostic request is issued, the EHR communicates with the architecture to identify relevant imaging studies stored in multiple PACS systems. The application queries each PACS using the DICOM standard to identify and retrieve imaging objects associated with the patient. These objects are then referenced and managed using the IHE XDS-I profile, which ensures image accessibility within the EHR ecosystem. Once retrieved, the imaging data are processed by the DIAM. This module performs automated preprocessing, lesion detection, and feature extraction.

All information, including clinical, anamnesis, and imaging data, is aggregated and validated using the HFFM. The CIIPM then calculates health status indicators and risk assessment metrics to support decision-making processes. All FHIR data resources are registered in the HIS using the Register Document Set-B [ITI-42] IHE transaction to ensure full interoperability and traceability across institutional repositories.

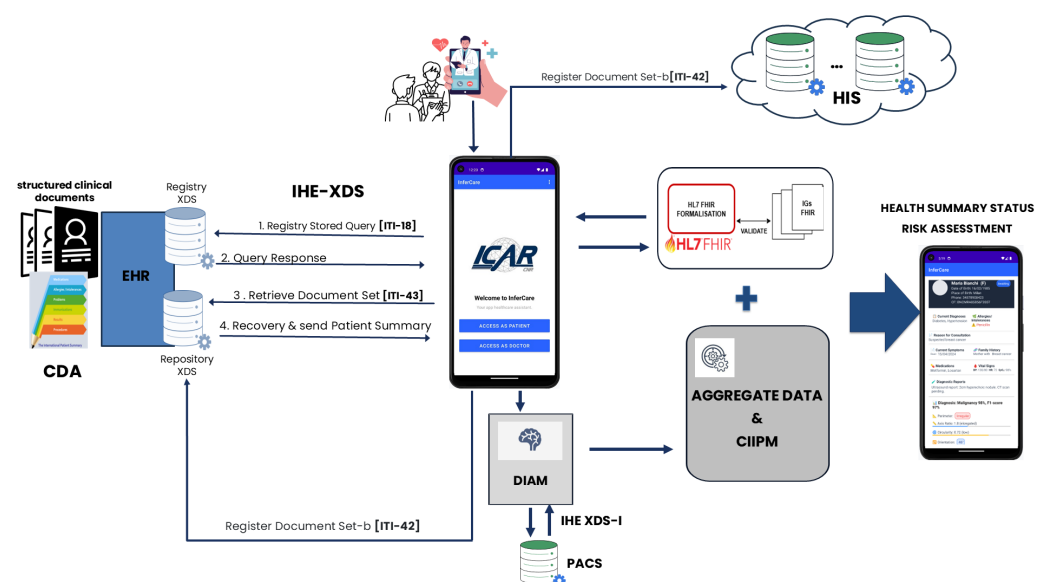


Figure 9. This diagram illustrates the data flow within the InferCare application and the interaction between heterogeneous medical data and external systems.

6. Conclusions

In this work, we utilized the HL7 FHIR standard [43] and all the standard tools [44,45] provided by the HL7 community to develop IPDSS, an architecture that enables the integration of data from diagnostic images, patient-collected data, and data from heterogeneous sources. We chose a case study that identifies AI-based risk assessment related to a clinical condition of breast cancer based on the analysis of US images by using ML models.

To manage the exchange of heterogeneous data and support multiple interactions with external modules and systems, the InferCare Android application was developed and seamlessly integrated within the proposed architecture. By leveraging the HL7 FHIR standard for data representation and communication, the application ensures interoperability across disparate HIS while providing an interactive graphical interface that facilitates diagnostic workflows and enhances user interaction for both patients and clinicians.

It also allows patients to record their medical history and physicians to obtain a summary of the most important information for an effective and rapid diagnosis, enabling risk assessments associated with the patient's clinical condition and other relevant information to be transformed into standard resources to be sent to external systems.

Furthermore, the goal of the architecture is to obtain useful information from different sources (EHR, PACS, medical history, etc.), standardize it, and make it available to all healthcare professionals who will be treating the patient (through EHR, HIS, etc. systems).

Thus, in the future, we intend to extend our work toward a more comprehensive and generalizable adoption of the proposed architecture, enabling the integration of unstructured clinical documents (from HIS, EHR, etc.), image-derived features across different diagnostic modalities, and information collected through patient-facing applications. Further research will focus on evaluating and validating the clinical impact of multimodal data fusion within the defined framework, assessing its contribution to diagnostic accuracy, optimizing clinical workflow, and providing personalized decision support. Moreover, future investigations could explore the incorporation of federated learning and privacy-preserving mechanisms to enable distributed AI model training across institutions without compromising data confidentiality. The integration of explainable AI components and continuous learning strategies will also be considered to enhance the transparency, adaptability, and long-term reliability of the proposed system in real-world clinical environments.

Author Contributions: Conceptualization, N.B. and M.S.; methodology, N.B. and M.S.; software, T.C. and M.R.; validation, N.B., M.S., T.C., M.R. and S.D.P.; formal analysis, N.B. and M.S.; investigation, T.C. and M.R.; resources, T.C., M.R. and S.D.P.; data curation, T.C., M.R. and S.D.P.; writing—original draft preparation, N.B. and M.S.; writing—review and editing, N.B., M.S., T.C., M.R. and S.D.P.; supervision, N.B. and M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Project “RIGOLETTO: Creation of intelligent management platform for oncology patients”—Bando “Accordi per l’innovazione”, del Ministero delle Imprese e del Made in Italy, mimit.AOO_IAI.REGISTRO INTERNO.R.0001470.08-05-2023, for providing support for this study.

Institutional Review Board Statement: This study does not involve human participants, human data, or human biological samples. The manuscript presents a theoretical framework and a conceptual architecture for integrating and processing health data within an FHIR-based system. Although examples referring to patient data collection are discussed, these are purely illustrative, and no real patient data were collected, accessed, or analyzed. Therefore, approval from an Institutional Review Board or Ethics Committee was not required.

Informed Consent Statement: Not applicable. This study does not involve human participants or the use of identifiable or non-identifiable human data.

Data Availability Statement: The image dataset BUSI, used for the experiments, is publicly available at <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset> (accessed on 19 March 2025). The Test IG “Remote Anamnesis” is publicly available at <http://remote-anamnesis.na.icar.cnr.it> (accessed on 8 August 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AI	Artificial Intelligence
ABUS	Automated Breast Ultrasound System
API	Application Programming Interface
CAD	Computer-Aided Diagnosis
DSS	Clinical Decision Support System
CIIPM	Clinical Information of Interest Presentation Module
CIC	Clinical Information Council
CIMI	Clinical Information Modeling Initiative
DL	Deep Learning
DIAM	Diagnostic Image Analysis Module
DT	Decision Tree
EHR	Electronic Health Record
ECG	Electrocardiogram
EMR	Electronic Medical Record
FHIR	Fast Healthcare Interoperability Resources
FR	Functional Requirement
HFFM	HL7 FHIR Formalization Module
HF	Hand-crafted Features
HIS	Health Information System
HL7	Health Level Seven International
ICE	Integrated Clinical Environment
ICHOM	International Consortium for Health Outcomes Measurement
IG	Implementation Guide
IHE	Integrating the Healthcare Enterprise
IPDSS	Integrated Patient Decision Support System
IPS	International Patient Summary
JSON	JavaScript Object Notation
LR	Logistic Regression
mCODE	Minimal Common Oncology Data Elements
ML	Machine Learning
MLP	Multi-Layer Perceptron
MDIRA	Medical Device Interoperability Reference Architecture
mHealth	Mobile Health
NB	Naive Bayes
OSIRIS	Open Standards for Interoperable and Reusable Information in Oncology
PF	Pathological Features
RF	Random Forest
RFE	Recursive Feature Elimination
REST	Representational State Transfer
SMOTE	Synthetic Minority Oversampling Technique
US	Ultrasound
XML	Extensible Markup Language
XDS	Cross-Enterprise Document Sharing

References

- Adler-Milstein, J.; Holmgren, A.J.; Kralovec, P. Electronic Health Record Adoption and Interoperability among U.S. Hospitals. *Health Aff.* **2021**, *40*, 1287–1296.
- Lentini, S.; Grosso, E.; Masala, G.L. A Comparison of Data Fragmentation Techniques in Cloud Servers. In *Proceedings of the Advances in Internet, Data & Web Technologies*; Barolli, L., Xhafa, F., Javaid, N., Spaho, E., Kolici, V., Eds.; Springer: Cham, Switzerland, 2018; pp. 560–571.
- Mercy, W.; Annabel, L.S.P. Secure Electronic Health Record Storage in the Cloud Based on Multiple Fragmentation and Reconstruction Using Blockchain with Cryptographic Techniques. In *Proceedings of the 2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)*, Gobichettipalayam, India, 12–13 December 2024; pp. 1232–1236. [\[CrossRef\]](#)
- Sreejith, R.; Senthil, S. Smart Contract Authentication assisted GraphMap-Based HL7 FHIR architecture for interoperable e-healthcare system. *Heliyon* **2023**, *9*, e15180. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ramgopal, S.; Sanchez-Pinto, L.N.; Horvat, C.M.; Carroll, M.S.; Luo, Y.; Florin, T.A. Artificial intelligence-based clinical decision support in pediatrics. *Pediatr. Res.* **2023**, *93*, 334–341. [\[CrossRef\]](#) [\[PubMed\]](#)
- Duda, S.N.; Kennedy, N.; Conway, D.; Cheng, A.C.; Nguyen, V.; Zayas-Cabán, T.; Harris, P.A. HL7 FHIR-based tools and initiatives to support clinical research: A scoping review. *J. Am. Med. Inform. Assoc.* **2022**, *29*, 1642–1653. [\[CrossRef\]](#)
- Taber, P.; Radloff, C.; Del Fiore, G.; Staes, C.; Kawamoto, K. New standards for clinical decision support: A survey of the state of implementation. *Yearb. Med Inform.* **2021**, *30*, 159–171. [\[CrossRef\]](#)
- Zhang, X.; Lin, X.; Tan, Y.; Zhu, Y.; Wang, H.; Feng, R.; Tang, G.; Zhou, X.; Li, A.; Qiao, Y. A multicenter hospital-based diagnosis study of automated breast ultrasound system in detecting breast cancer among Chinese women. *Chin. J. Cancer Res.* **2018**, *30*, 231. [\[CrossRef\]](#)
- Moon, W.K.; Lee, Y.W.; Ke, H.H.; Lee, S.H.; Huang, C.S.; Chang, R.F. Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. *Comput. Methods Programs Biomed.* **2020**, *190*, 105361. [\[CrossRef\]](#)
- Badawy, S.M.; Mohamed, A.E.N.A.; Hefnawy, A.A.; Zidan, H.E.; GadAllah, M.T.; El-Banby, G.M. Classification of breast ultrasound images based on convolutional neural networks—a comparative study. In *Proceedings of the 2021 International Telecommunications Conference (ITC-Egypt)*, Alexandria, Egypt, 13–15 July 2021; pp. 1–8.
- Seiler, S.J.; Neuschler, E.I.; Butler, R.S.; Lavin, P.T.; Dogan, B.E. Optoacoustic imaging with decision support for differentiation of benign and malignant breast masses: A 15-reader retrospective study. *Am. J. Roentgenol.* **2023**, *220*, 646–658. [\[CrossRef\]](#)
- Kwon, H.; Oh, S.H.; Kim, M.G.; Kim, Y.; Jung, G.; Lee, H.J.; Kim, S.Y.; Bae, H.M. Enhancing Breast Cancer Detection through Advanced AI-Driven Ultrasound Technology: A Comprehensive Evaluation of Vis-BUS. *Diagnostics* **2024**, *14*, 1867. [\[CrossRef\]](#)
- Wang, S.; Zhao, Z.; Ouyang, X.; Liu, T.; Wang, Q.; Shen, D. Interactive computer-aided diagnosis on medical image using large language models. *Commun. Eng.* **2024**, *3*, 133. [\[CrossRef\]](#)
- Azam, S.; Montaha, S.; Raiaan, M.A.K.; Rafid, A.R.H.; Mukta, S.H.; Jonkman, M. An automated decision support system to analyze malignancy patterns of breast masses employing medically relevant features of ultrasound images. *J. Imaging Inform. Med.* **2024**, *37*, 45–59. [\[CrossRef\]](#)
- Ragab, M.; Albukhari, A.; Alyami, J.; Mansour, R.F. Ensemble deep-learning-enabled clinical decision support system for breast cancer diagnosis and classification on ultrasound images. *Biology* **2022**, *11*, 439. [\[CrossRef\]](#) [\[PubMed\]](#)
- Daoud, M.I.; Abdel-Rahman, S.; Bdair, T.M.; Al-Najar, M.S.; Al-Hawari, F.H.; Alazrai, R. Breast tumor classification in ultrasound images using combined deep and handcrafted features. *Sensors* **2020**, *20*, 6838. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cruz-Ramos, C.; Garcia-Avila, O.; Almaraz-Damian, J.A.; Ponomaryov, V.; Reyes-Reyes, R.; Sadovnychiy, S. Benign and malignant breast tumor classification in ultrasound and mammography images via fusion of deep learning and handcraft features. *Entropy* **2023**, *25*, 991. [\[CrossRef\]](#)
- Khan, M.U.; Bianconi, F.; Du, H.; Jassim, S. Hand-crafted Vs. deep CNN features to distinguish benign from malignant lesions in breast ultrasound images. In *Proceedings of the 2025 International Conference on Control, Automation and Diagnosis (ICCAD)*, Barcelona, Spain, 1–3 July 2025; pp. 1–6.
- Abhisheka, B.; Biswas, S.K.; Purkayastha, B.; Das, S. Integrating deep and handcrafted features for enhanced decision-making assistance in breast cancer diagnosis on ultrasound images. *Multimed. Tools Appl.* **2025**, *84*, 43263–43285. [\[CrossRef\]](#)
- Taheri, F.; Rahbar, K. Improving breast cancer classification in fine-grain ultrasound images through feature discrimination and a transfer learning approach. *Biomed. Signal Process. Control* **2025**, *106*, 107690. [\[CrossRef\]](#)
- Ellis, J.; Appiah, K.; Amankwaa-Frempong, E.; Kwok, S.C. Classification of 2d ultrasound breast cancer images with deep learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 17–18 June 2024; pp. 5167–5173.
- Mandel, J.C.; Kreda, D.A.; Mandl, K.D.; Kohane, I.S.; Ramoni, R.B. SMART on FHIR: A standards-based, interoperable apps platform for electronic health records. *J. Am. Med. Inform. Assoc.* **2016**, *23*, 899–908. [\[CrossRef\]](#)

23. Eapen, B.R.; Archer, N.; Sartipi, K.; Yuan, Y. Drishti: A Sense-Plan-Act Extension to Open mHealth Framework Using FHIR. In Proceedings of the 2019 IEEE/ACM 1st International Workshop on Software Engineering for Healthcare (SEH), Montreal, QC, Canada, 27 May 2019; pp. 49–52.
24. Sloane, E.B.; Cooper, T.; Silva, R. MDIRA: IEEE, IHE, and FHIR Clinical Device and Information Technology Interoperability Standards, bridging Home to Hospital to “Hospital-in-Home”. In Proceedings of the SoutheastCon 2021, Virtual, 10–14 March 2021; pp. 1–4.
25. Lee, J.; Kim, J. Design of an ECG Stream Analysis Framework Based on FHIR Data Model. In Proceedings of the 2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN), Budapest, Hungary, 2–5 July 2024; pp. 567–569.
26. Osterman, T.J.; Terry, M.; Miller, R.S. Improving Cancer Data Interoperability: The Promise of the Minimal Common Oncology Data Elements (mCODE) Initiative. *JCO Clin. Cancer Inform.* **2020**, *4*, 993–1001. [\[CrossRef\]](#)
27. Chen, J.; Chiang, Y. Applying the Minimal Common Oncology Data Elements (mCODE) to the Asia-Pacific Region. *JCO Clin. Cancer Inform.* **2021**, *5*, 252–253. [\[CrossRef\]](#)
28. Guérin, J.; Laizet, Y.; Le Texier, V.; Chanas, L.; Rance, B.; Koeppel, F.; Lion, F.; Gourgou, S.; Martin, A.L.; Tejeda, M.; et al. OSIRIS: A Minimum Data Set for Data Sharing and Interoperability in Oncology. *JCO Clin. Cancer Inform.* **2021**, *5*, 256–265. [\[CrossRef\]](#)
29. Major, V.J.; Wang, W.; Aphinyanaphongs, Y. Enabling AI-Augmented Clinical Workflows by Accessing Patient Data in Real-Time with FHIR. In Proceedings of the 2023 IEEE 11th International Conference on Healthcare Informatics (ICHI), Houston, TX, USA, 26–29 June 2023; pp. 531–533.
30. Chishtie, J.; Sapiro, N.; Wiebe, N.; Rabatach, L.; Lorenzetti, D.; Leung, A.A.; Rabi, D.; Quan, H.; Eastwood, C.A. Use of Epic electronic health record system for health care research: scoping review. *J. Med. Internet Res.* **2023**, *25*, e51003. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Peretokin, V.; Basdekis, I.; Kouris, I.; Maggesi, J.; Sicuranza, M.; Su, Q.; Acebes, A.; Bucur, A.; Mukkala, V.J.R.; Pozdniakov, K.; et al. Overview of the SMART-BEAR Technical Infrastructure. In Proceedings of the 8th International Conference on Information and Communication Technologies for Ageing Well and e-Health, Online, 23–25 April 2022; SciTePress: Setúbal, Portugal, 2022; pp. 117–125. [\[CrossRef\]](#)
32. Conte, T.; Sicuranza, M. *Sistema FHIR-Based per la Raccolta Strutturata dell’Anamnesi in Remoto: Approccio, Standard e Implementazione*; Rapporto Tecnico RT-ICAR-NA-2025-05; CNR-ICAR: Napoli, Italy, 2025.
33. HL7 International. International Patient Summary (IPS) Implementation Guide. Available online: <https://build.fhir.org/ig/HL7/fhir-ips/> (accessed on 25 June 2025).
34. Teresa Conte and Mario Sicuranza. Remote Anamnesis Implementation Guide for Technical Report. Available online: <https://anamnesi.na.icar.cnr.it/> (accessed on 8 August 2025).
35. Teresa Conte and Mario Sicuranza. Remote Anamnesis Implementation Guide. Available online: <http://remote-anamnesis.na.icar.cnr.it> (accessed on 8 August 2025).
36. Guo, R.; Lu, G.; Qin, B.; Fei, B. Ultrasound imaging technologies for breast cancer detection and management: A review. *Ultrasound Med. Biol.* **2018**, *44*, 37–70. [\[CrossRef\]](#)
37. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 26.
38. Magee, J.F. *Decision Trees for Decision Making*; Harvard Business Review: Brighton, MA, USA, 1964.
39. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. *Learning Internal Representations by Error Propagation*; Technical Report; Institute for Cognitive Science, California University San Diego: La Jolla, CA, USA, 1985.
40. Langley, P.; Iba, W.; Thompson, K. An analysis of Bayesian classifiers. In Proceedings of the AAAI, San Jose, CA, USA, 12–16 July 1992; Citeseer: University Park, PA, USA, 1992; Volume 90, pp. 223–228.
41. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp. 278–282.
42. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [\[CrossRef\]](#)
43. Standard HL7 FHIR. HL7 Documentation. Available online: <https://www.hl7.org/fhir/documentation.html> (accessed on 15 May 2025).
44. HL7. SUSHI – SUSHI Unshortens Short Hand Inputs. Available online: <https://build.fhir.org/ig/HL7/fhir-shorthand/overview.html> (accessed on 15 May 2025).
45. HL7. IG Publisher. Available online: <https://confluence.hl7.org/plugins/servlet/mobile?contentId=35718627#content/view/35718627> (accessed on 15 May 2025).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.