ICT ECOSYSTEMS FOR PRECISION MEDICINE

According to the Precision Medicine Initiative, precision medicine is an "emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person". Precision Medicine allows treating doctors to decide on the appropriate treatment for a patient who is suffering from a specific disease and who has a specific genotype. The Precision Medicine approach differs from the traditional approach in which the treatment of diseases follows a "one-size-fits-all" approach focusing on the average person without considering the differences exhibited by the various patients. Precision medicine is a model in which the individuals' phenotype and genes, environment, and lifestyle are used as additional layers of patient data in disease treatment and prevention plans. All Precision Medicine strategies include the use of decision-making processes based on the use of biomarkers. Biomarkers can be used to augment and refine traditional patient data to provide a higher level of specificity for disease prevention and patient treatment and are crucial in providing the appropriate treatment to her.
In Greece currently there are three Precision Medicine initiatives:

- The Hellenic Precision Medicine Network on Cancer (ONCOPMNET) (May 2018 – December 2021)
- The National Network For Genetic Cardiovascular Diseases Study And Prevention Of Sudden Death In The Young On The Basis Of Precision Medicine (GR iCARDIACNET) (July 2019 – July 2022) and
- The National Research Network for Genetic Neurodegenerative Diseases (GR NEUROPMNET) (May 2020 – May 2022)

The Institute of Computer Science-FORTH, participates in all the aforementioned Networks, and is the only ICT participant for GR iCARDIACNET and GR NEUROPMNET. Irini Fundulaki is the scientific lead on behalf of the Institute of Computer Science - FORTH for all the aforementioned Precision Medicine Networks.

The mission of the Hellenic Precision Medicine Network on Cancer is to (a) connect the Network with the National Health System (b) provide high-quality healthcare to Greek citizens (c) enrich diagnosis knowledge and prediction outcome and (d) improve the targeted therapeutic treatment of cancer patients. The ICT objectives of the Network are the development of custom bioinformatics pipelines for different types of cancers.

The National Network for Genetic Cardiovascular Diseases Study And Prevention Of Sudden Death In The Young On The Basis Of Precision Medicine (GR iCARDIACNET), aims at the promotion of Precision Medicine in cardiology and the prevention of sudden death in the young including the study of hereditary cardiovascular diseases in Greece. The ICT objectives of the Network include the creation of Registries that store clinical and genomic patient data for inherited cardiovascular diseases, and forensic reports. The ultimate goal is to use the data stored in the registries to find populations who carry a specific genetic mutation and break the chain of the inherited disease.

The mission of the National Initiative for Neurodegenerative Disease Research Based on Precision Medicine (GR DeGeNEURO NET) is the promotion of Precision Medicine for Neurodegenerative Diseases, with emphasis on their genetic basis. Similar to GR iCARDIACNET, the ICT objectives of the Network include the creation of Registries that store clinical and genomic patient data for neurodegenerative diseases.

For GR iCARDIACNET, we developed an innovative ICT Ecosystem [FSP+20] that consists of a set of systems that support clinical doctors and the geneticists who manage clinical and genomic patient data respectively, and the epidemiologists who want to be aware of the epidemics of a disease to decide on actions that must be undertaken in order to address the challenges presented. The design of all ICT systems is based on requirements provided by the aforementioned agents and on results of discussions with the Network's Data Protection Officer

and Legal Team and follows the FAIR principles: all data are Findable, Accessible, Interoperable and Reusable, for both machines and people:

Findable: We follow a patient-centric design (Figure 1) in which each patient is uniquely identified by her social security number and assigned a unique system identifier used for the complete data life cycle. The patient-centric approach supports the integration of diverse data (clinical, genomic, imaging data).

Accessible: Data and metadata stored in the ICT Ecosystem are accessible using standard protocols which are open, free and universally implementable. Furthermore, due to the sensitive nature of the data, we designed and developed a role-based access control framework in which we consider different roles and groups with different access rights that determine who (role or group) has a certain type of access (read, write, modify) for specific data (patient record or parts thereof). We must note that an appropriate authentication mechanism is in place in order to ensure that only authorized users have access to the data. Last but not least, patient's data is anonymized and encrypted when exchanged between systems to ensure that sensitive patient information is not revealed to unauthorized users.

Interoperable: We use state-of-the-art schemas for both data and metadata that ensure that data is exchanged between the different systems of the ICT Ecosystem but also with external systems such as National Registries, the National e-prescription system, Hospital Information Systems among others.

Reusable: Both data and metadata are released with a clear and accessible data usage license and are accompanied with necessary provenance information to ensure accountability.

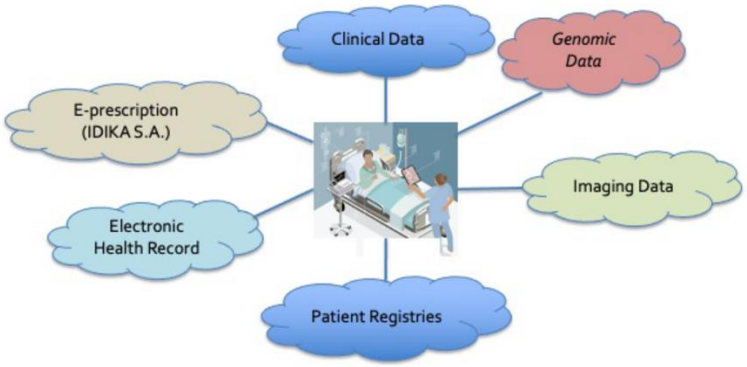

Figure 1: Patient-centric design

The ICT Ecosystem consists of the following systems:

- **Inherited Cardiovascular Diseases Registry** used by clinical doctors to store patient information organized by visit (symptoms, diagnosis, comorbidities, medications, investigations, procedures among others) allowing the doctor to record the evolution of the patient's health. The schema registry is based on the Clinical protocol of the European Society of Cardiology / EURObservational Research Programme that focuses on Cardiomyopathies and Myocarditis. The Registry interoperates with the National Social Security Registry (ATLAS) of IDIKA S.A. to retrieve updated patient information and supports the secure storage, and controlled access to sensitive patient health data.
- **Sudden Death in the Young Registry** used by pathologists to record forensic reports. The schema of the registry is based upon the requirements of the pathologists and cardiologists who participate in GR iCARDIACNET.
- **Genetic Data Repository** used by the lab personnel to store the analysis files and the associated metadata (variant, classification, exome, zygocity). We developed an application that produces in a semi-automatic way the genetic report that is sent to the clinician who orders the genetic test.
- **e-genrefappr** system used by the members of the Genetic Referral Approval Committee to approve or disapprove the requests of the clinicians for genetic tests. The anonymized clinical data of the patient are made available to the members of the committee who can then accept or reject the clinician's request.

- **e-genreferral** system that enables clinicians to prescribe genetic tests to patients. The system retrieves the anonymized patient's clinical data that are necessary for the lab geneticists to perform the genetic analysis. Each referral has a unique code used to mark the patient's sample sent to the lab for analysis.
- **e-labBook** system, subsystem of e-genreferral that allows the clinician to follow the progress of the analysis of the patient's sample and supports the lab personnel in their everyday practice. The sample recipient can use the system to record the sample's arrival date and whether it is appropriate for analysis, the technical supervisor to assign the analysis to a geneticist, the scientific supervisor to sign the genetic report and upload the analysis files and the genetic report to the repository. The system follows the ISO15189 standard that specifies the quality management system requirements particular to medical laboratories and is use case independent (i.e., can be used by any laboratory that adheres to the ISO 15189 restrictions).

The ICT ecosystem supports an information flow that is entirely digitized (Figure 2). In short, a clinician can search for the patient in the Registry of Inherited Cardiovascular Diseases using the patient's Social Security Number. A request is sent to the ATLAS system of IDIKA S.A to obtain the patient's demographic data (i.e., first, last name, age, contact details). Once the patient's record is created in the Registry, the clinician can then add clinical data to it. The clinician can then prescribe (if sufficient information is entered in the registry) genetic tests using the e-genreferral system. A subset of the patient's anonymized clinical data is sent from the Registry to e-genreferral and subsequently the referral request is sent to the e-genrefappr system. The members of the Genetic Referral Approval Committee examine the clinician's request and if the request is accepted, the anonymized patient information is sent to the laboratory that will perform the sample analysis. The e-LabBook system is used by the lab personnel as described above to manage the patient sample and proceed to the analysis of the sample. At the end of the analysis the laboratory personnel can store the primary and secondary data of the analysis in the Genetic Data Repository as well as the genetic report that is also delivered to the clinician who prescribed the genetic tests.

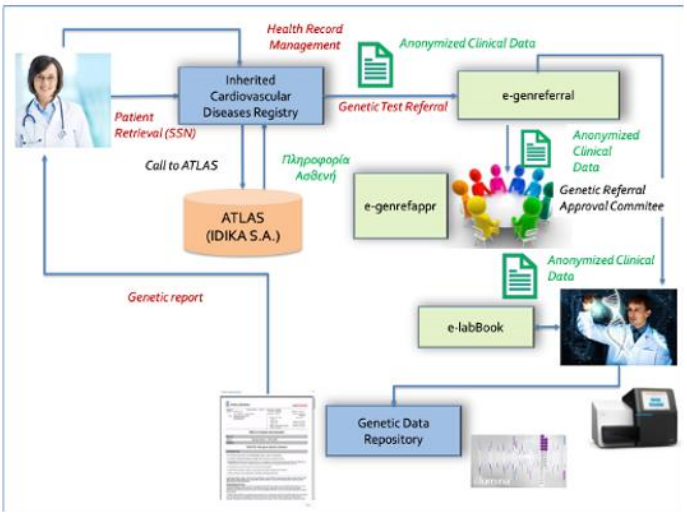

Figure 2: Information Flow

The ICT Ecosystem is used by the clinicians and the lab personnel of hospitals and research centers that participate in the Network extensively to store the clinical and genomic data of the patients. More specifically, the ICT Ecosystem is used in the everyday practice of medical doctors from 6 large hospitals in Greece (Onassis Cardiac Surgery Center, Evaggelismos General Hospital, Ippokrateio General Hospital of Athens, Ippokrateio - General Hospital of Thessaloniki, AHEPA Hospital, and University General Hospital of Heraklion), from 2 Analysis Laboratories (Onassis Cardiac Surgery Center, Institute of Molecular Biology and Biotechnology – Foundation for Research and Technology) and last by the pathologists of the Laboratories of Forensic Medicine & Toxicology of the National and Kapodistrian University of Athens and the Aristotle University of Thessaloniki).

The data stored in the systems are used by a special purpose application to extract the epidemiological map of the country for the diagnosis and the variants that are associated with patients in the Network.
One year after the operation of the Ecosystem, **1599 patient records** are stored in the **Inherited Cardiovascular Diseases Registry, 67 forensic reports** are added in the **Sudden Death in the Young Registry, 239 genetic reports** are stored in the **Genetic Data Repository, 532 genetic tests** were prescribed using the **e-genreferral system** and finally, **298 biological samples** were processed using the **e-labBook system.**

The data stored in the Ecosystem is used to produce a set of interesting statistical results that provide a picture of the health status of the population regarding cardiovascular diseases. Figure 3 provides the distribution of different types of diagnosis, Figure 4 presents the cases of cardiomyopathies (a) and myocarditis (b) in Greece, Figure 5 shows the pathogenic and likely pathogenic variants found in patients who are diagnosed with cardiomyopathy and last Figure 6 presents the pathogenic and likely pathogenic variants for the different diseases.

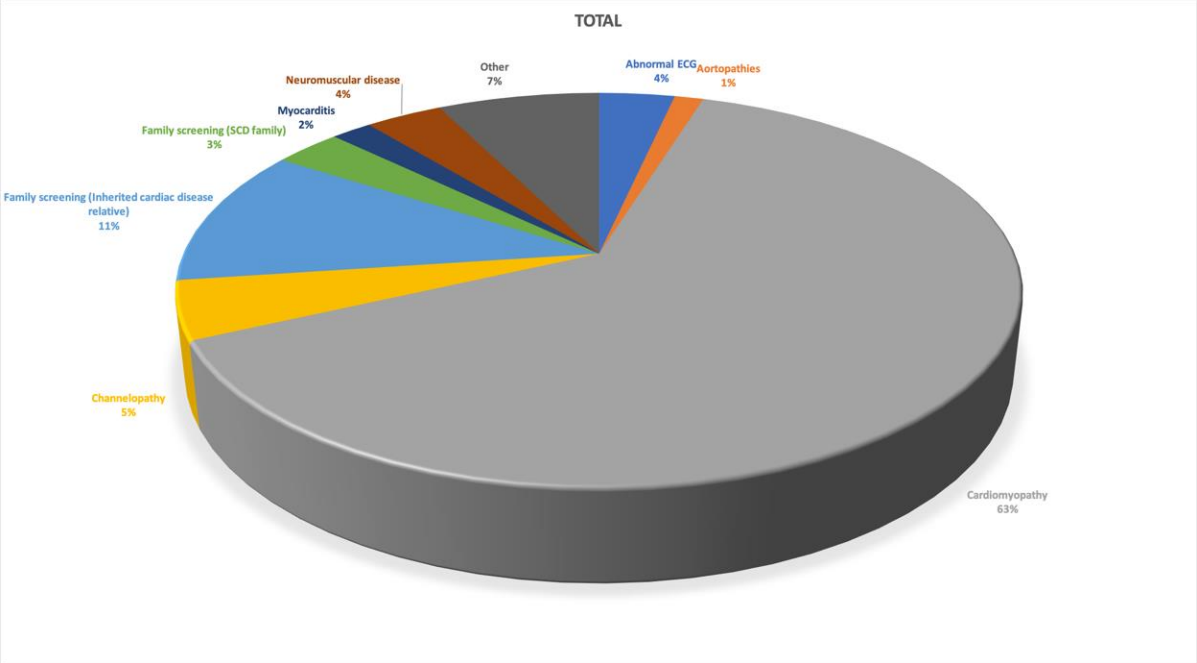

Figure 3: Distribution of different types of diagnosis

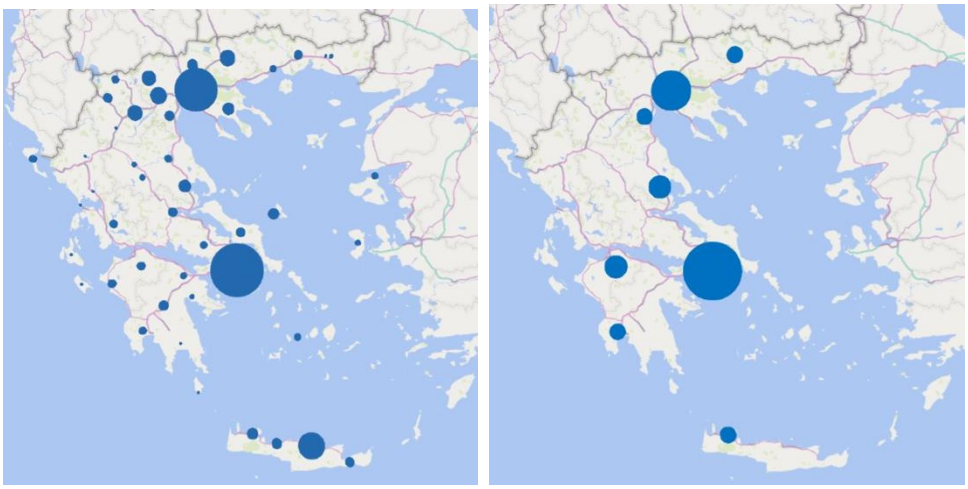

Figure 4: Occurrence of Cardiomyopathies (left), Myocarditis (right) in Greece

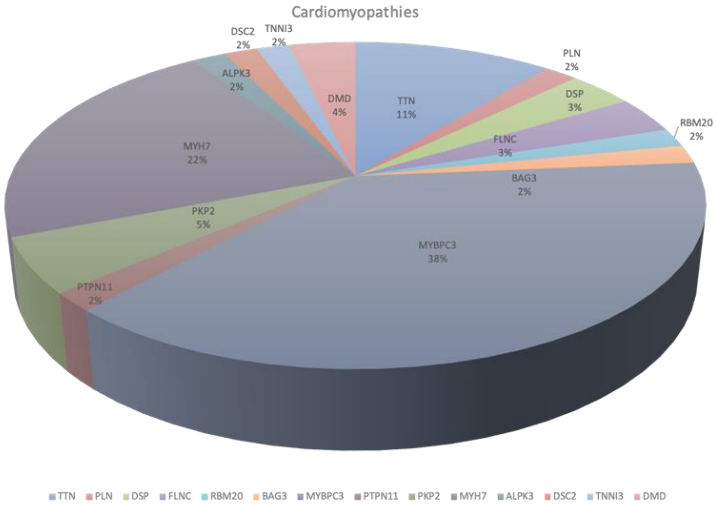

Figure 5: Variants associated with Cardiomyopathies

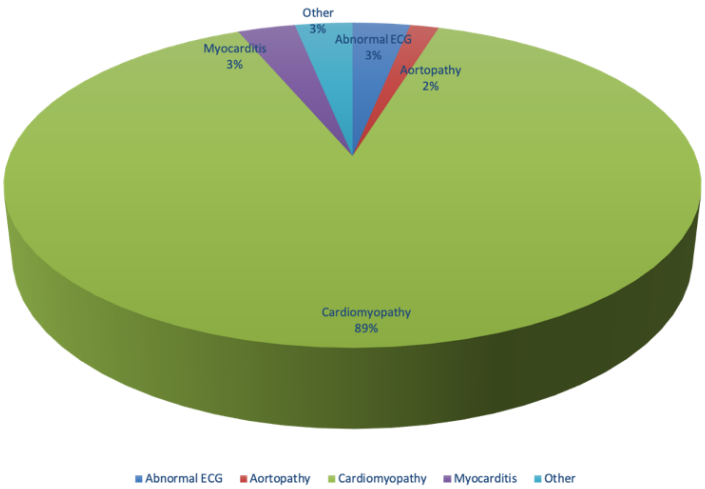

Figure 6: Pathogenic and Likely Pathogenic Variants (NGS analysis)

For the ICT Ecosystem we followed a microservices architecture and for the implementation of the aforementioned systems we used virtualization and containerization technologies. Drupal v8 Content Management System (CMS) was used as the front-end, and for the User Interface we employed state of the art technologies such as HTML, CSS, and JS framework Bootstrap 3, jQuery, Apache Web Server, PHP. Data is stored in the relational database system MySQL. We used Keycloak as the central user management system for all

systems that support Single sign-on (SSO), Single log-out (SLO). To secure data we used encryption, and the front-end back-end communication is performed via an encrypted communication channel.

In the context of GR NEUROPMNET we are developing a Registry for the following neurodegenerative diseases: Parkinson's, Huntington, Rapid Eye Movement Sleep Behavior Disorder (RBD), Multiple Sclerosis, Amyotrophic Lateral Sclerosis (ALS), Alzheimer. The challenge in this case was to identify common information for the diseases so that we could implement a registry that a medical doctor could add clinical data for patients diagnosed for different diseases. The ICT Ecosystem developed in the context of GR iCARDIACNET will be adapted to address the requirements of GR NEUROPMNET that mostly pertain to the type of genetic tests (Whole Exome Sequencing instead of specific sets of genes – panels – examined currently in the context of GR iCARDIACNET).

For ONCOPMNET we developed the first version of the e-genreferral system used by all the oncologists in Greece to prescribe genetic tests to cancer patients. Users are authenticated using the authentication mechanism of the National e-prescription system of IDIKA S.A.