CrowdHEALTH

Collective Wisdom Driving Public Health Policies

Del. no. – D6.7 Use Case Scenarios Definition and Design

Project Deliverable

This project has received funding from the European Union's Horizon 2020 Programme (H2020-SC1-2016-CNECT) under Grant Agreement No. 727560

D6.7 Use Case Scenarios Definition and Design

Work Package:		WP6
Due Date:		31/10/2017
Submission Date:		21/11/2017
Start Date of Project:		01/03/2017
Duration of Project:		36 Months
Partner Responsible of Deliverable:		Karolinska Institutet
Version:		2.3
Status:	☒ Final ☐ Draft ☐ Ready for internal Review ☒ Task Leader Accepted ☒ WP leader accepted ☒ Project Coordinator accepted	
Author name(s):	Sokratis Nifakos (KI) Tanja Tomson (KI) Jan Jannsen (DFKI) Salvador Tortajada Velert (HULAFE) Thanos Kosmidis (CRA) Christos Panagopoulos (BIO) Gregor Jurak (ULJ) Marjeta Kovač (ULJ) Gregor Starc (ULJ	
Reviewer(s):	Marta Patiño (UPM)	Vergard Egen (IT-INN) Daniel Burns (IT-INN)
Nature:	☒ R – Report ☐ D – Demonstrator	
Dissemination level:	☒ PU – Public ☐ CO – Confidential ☐ RE – Restricted	

REVISION HISTORY			
Version	Date	Author(s)	Changes made
0.1	23/08/2017	Sokratis Nifakos	TOC to share with the participants.
0.2	31/08/2017	Tanja Tomson	Review
0.3	05/10/2017	Sokratis Nifakos	Introduction
0.4	12/10/2017	Sokratis Nifakos	Use Case KI
0.5	16/10/2017	Sokratis Nifakos	Integrated information from UC partners
0.6	17/10/2017	Sokratis Nifakos	Conclusion and further work
1.1	27/10/2017	Salvador Tortajada	New version of HULAFE Use Case
1.2	31/10/2017	Christos Panagopoulos	BIO part
1.3	31/10/2017	Gregor Jurak	ULJ part
1.4	31/10/2017	Thanos Kosmidis	CRA part
1.5	31/10/2017	Sokratis Nifakos	KI part and Introduction
1.6	09/11/2017	Christos Panagopoulos	BIO use case updated
1.7	09/11/2017	Gregor Jurak	ULJ use case updated
1.8	09/11/2017	Thanos Kosmidis	CRA use case updated
1.9	10/11/2017	Sokratis Nifakos	Updated all sections based on reviewers' comments. Introduction, Executive Summary, Conclusion. Finalized document (formatting).
2.0	15/11/2017	Salvador Tortajada	Updated HULAFE Use Case
2.1	15/11/2017	Jan Janssen	Updated DFKI-HULAFE and DFKI-CRA parts
2.2	15/11/2017	Sokratis Nifakos	Merged all information.
2.3	21/11/2017	ATOS	Quality review and submission to EC

List of acronyms

ADL	Activities of daily living
ATC	Anatomical Therapeutic Chemical Classification System
ASCII	American Standard Code for Information Interchange
BIO	BioAssist
BMI	Body Mass Index
BODE	An acronym for Body-mass index, airflow Obstruction, Dyspnea, and Exercise.
bpm	Beats per minute
CRA	CareAcross
COPD	Chronic Obstructive Pulmonary Disease
CO2	Carbon dioxide
CRP	C-reactive protein (CRP) is an annular (ring-shaped), pentameric protein found in blood plasma, whose levels rise in response to inflammation.
CSV	Comma separated values file
CVD	Cardiovascular disease
DFKI	Deutsche Forschungszentrum für Künstliche Intelligenz
dL	deciliter
EHR	Electronic Health Record
FEV 1	Forced expiratory volume in 1 second
FSUL	Faculty of Sport, University of Ljubljana
GOT	Glutamyl oxaloacetic transaminase
GPS	Global Positioning System
GPT	Glutamyl pyruvic transaminase
HHR	Holistic Health Record
HULAFE	Fundaction para la Investigacion del Hospital Universitario La Fe de la Comunidad Valenciana
ICD-10	International Statistical Classification of Diseases and Related Health Problems 10th Revision.
IPF	Idiopathic pulmonary fibrosis
IRIS	Information System for anatomical pathology, laboratories and x-ray departments in HULAFE
kg	Kilogram
KI	Karolinska Institutet
KPI	Key Performance Indicator
LDL	Low Density Cholesterol
L/min	Liter per minute
mg	Miligrams
NAOS	Spanish Strategy for nutrition, physical activity and prevention of obesity
NCDs	Noncommunicable diseases
ORION	Orion Clinic is the system that provides instant access to real time electronic patient care records to healthcare professionals in HULAFE
O2	Oxigene
pH	Potential of Hydrogen, a scale of acidity from 0 to 14.
R	A free software environment for statistical computers and graphics
RIKS HIA	Swedish Register of Cardiac Intensive Care

SCAAR	Swedish Coronary Angiography and Angioplasty Register
SD	Standard deviation
SEPHIA	Swedish Registry for secondary prevention.
SES	Socio-economic status
SIA-ABUCASIS	System of ambulatory Information in Valencia (Spain) with pharmacy prescription module and agenda of visits to the doctor
SLOFit	Slovenian surveillance system for physical and motor development of children and youth which was formerly known as Sports Educational Chart.
SpO2	Peripheral capillary oxygen saturation
SQL	Structured Query Language
TSH	Thyroid-Stimulating Hormone
T3	Triiodothyronin
T4	Thyroxine
ULJ	University of Ljubljana
UML	Unified Modeling Language
VAL	Vårdanalysdatabasen; Stockholm regional health care data warehouse
WHO	World Health Organization
WP	Work Package

Contents

1. Executive Summary 9
2. Introduction 10
3. Use Cases Definition and Design 11
3.1. HULAFE - Overweight and obesity 11
3.1.1. Brief Introduction 11
3.1.2. Key Players/Actors 12
3.1.3. Relevant Policies 13
3.1.4. Key Performance Indicators 13
3.1.5. Expected Outcomes 15
3.1.6. Data Flow 15
3.1.7. Contributions to HHRs 16
3.2. BioAssist – Chronic Disease Management 17
3.2.1. Brief Introduction 17
3.2.2. Key Players/Actors 18
3.2.3. Relevant Policies 19
3.2.4. Key Performance Indicators 19
3.2.5. Expected Outcomes 21
3.2.6. Data Flow 22
3.2.7. Contributions to HHRs 22
3.3. CareAcross-Coaching for cancer patients 25
3.3.1. Brief Introduction 25
3.3.2. Key Players/Actors 26
3.3.3. Relevant Policies 27
3.3.4. Key Performance Indicators 27
3.3.5. Expected Outcomes 28
3.3.6. Data Flow 28
3.3.7. Contributions to HHRs 29
3.4. University of Ljubljana-Physical fitness, physical activity and obesity 31
3.4.1. Brief Introduction 31
3.4.2. Key Players/Actors 32
3.4.3. Relevant Policies 33

3.4.4. Key Performance Indicators 33
3.4.5. Expected Outcomes 35
3.4.6. Data Flow 35
3.4.7. Contributions to HHRs 36
3.5. Karolinska Institutet 37
3.5.1. Brief Introduction 37
3.5.2. Key Players/Actors 38
3.5.3. Relevant Policies 39
3.5.4. Key Performance Indicators 39
3.5.5. Expected Outcomes 42
3.5.6. Data Flow 42
3.5.7. Contributions to HHRs 42
3.6. DFKI-HULAFE and DFKI-CRA Integrated Use Cases 43
3.6.1. Brief Introduction 43
3.6.2. Key Players/Actors 48
3.6.3. Relevant Policies 49
3.6.4. Key Performance Indicators 49
3.6.5. Expected Outcomes 50
3.6.6. Data Flow 50
3.6.7. Contributions to HHRs 50
4. Conclusions 52
5. References 53

List of figures

Figure 1: Data Flow HULAFE Use Case 16
Figure 2: CRA Data Flow 29
Figure 3: Data flow between the key players 36
Figure 4: Key players/Actors flow diagram 39
Figure 5: Schema illustrating the cycle of the pilot study of the integrated Use Cases 46

List of tables

Table 1 Table of BioAssist Key Players/Actors 18
Table 2: BioAssist KPIs 20
Table 3: Other BioAssist KPIs 20
Table 4: BioAssist Devices 23
Table 5: CRA inputs about HHR Nutrition 30
Table 6: University of Ljubljana KPIs 34

1. Executive Summary

This document presents deliverable D6.7 Use Case Scenarios Definition and Design v1 of Work Package 6, Use Cases Adaptation, Integration and Experimentation. The main objective is to provide the Use Case Scenarios definition and specification as well as to present the scenarios in conjunction with the identification of the involved stakeholders that are vital for the deployment of the CrowdHEALTH platform.

In this deliverable, we aim to describe the representative use case scenarios for the CrowdHEALTH project. Through these use case scenarios, we aim to integrate the research and development work of the project, and utilise the use cases for the purposes of verification and collecting feedback on the use of the envisaged CrowdHEALTH platform. In continuous, we are focusing on the expecting outcomes after CrowdHEALTH platform implementation and we present an overview of the services that the platform will provide to the stakeholders.

The use cases included in the CrowdHEALTH consist of both private and public stakeholder and with different health scope around Europe.

Supporting Use Cases:

- HULAFE - Overweight and Obesity
- Karolinska Institutet - Cardiovascular Diseases
- ULJ - Physical Fitness, Physical Activity and Obesity
- CareAcross - Coaching for Cancer Patients
- BioAssist - Chronic Disease Management
- DFKI-HULAFE - Integrated Use Case
- DFKI-CRA – Integrated Use Case

Among the use cases, in this document we present also a synergy between three partners in CrowdHEALTH project: DFKI, HULAFE and CRA. The synergy aims to explore how DFKI could provide technical equipment for HULAFE and CRA in order to collect more information from patients with overweight and obesity and enhance further the holistic health records in this case.

2. Introduction

WP6 focuses on integrating the outcomes of the research and development WPs in CrowdHEALTH, verifying their applicability through five representative and two integrated use case scenarios use providing useful feedback about the CrowdHEALTH concepts and technologies.

Use Case Definition and Design describes the current status of requirements analysis from the Use Case partners and translates how users in the use cases perceive a realistic situation in a structural description. The cornerstone for the requirements elicitation and the scenarios definition has been the early engagement of all the project stakeholders into the scenarios execution. The deliverable highlights the initial inputs on pilots and technical deals concerning the components definition based on the Holistic Health Record (HHR) analysis and pilot preparation as part of the overall CrowdHEALTH scenarios execution.

Moreover, the Use Case scenarios provide insights by describing the respective typical situations that occur at the partners' sites. They provide critical information for the technical infrastructure of the project since they include distinct flows of actions or different contextual situations. Based on the individual nature and characteristics of each scenario, an analysis of HHR data that will be generated from the CrowdHEALTH platform has been carried on despite this is not strictly required during these early phases of the project.

Additionally, the document describes the key performance indicators for health policy creation, evaluation and implementation in the different use case sites. These performance indicators are important for the comprehension of the existing health policies and the ways that can be evaluated and improved in addition to create and implement new ones.

Finally, in the last section a description of synergy between different stakeholders in the project has been conducted. The description includes the phases of integration and the processes that the stakeholders will follow in correlation with CrowdHEALTH platform features.

3. Use Cases Definition and Design

3.1. HULAFE - Overweight and obesity

3.1.1. Brief Introduction

The World Health Organization (WHO) defines overweight and obesity as “abnormal or excessive fat accumulation that may impair health” [1]. This is currently measured using the Body Mass Index (BMI), which is a useful population-level measure. BMI is defined as weight/(height)2, where the weight is measured in kilograms and the height in meters. Hence, the BMI is measured in (kg/m2) units. However, BMI is considered a rough guide because it may not correspond to the same degree of fatness in different individuals.

The increasing concern about the prevalence of obesity and overweight is due to the association with the main chronic disease of our time: cardiac diseases, diabetes mellitus type II, hypertension, and some types of cancer. A higher BMI is correlated with higher levels of comorbidities and mortality due to such chronic diseases. In adults, obesity is also related with osteoarthritis and respiratory diseases. Obesity and overweight are harmful to health, both by themselves, as well as being risk factors for other chronic diseases and also as shortening factors for life expectancy, where changes in diet and sedentary lifestyles are the main triggers in the increase of obesity. All these conditions make overweight and obesity a main problem for public health being already considered by some experts as an epidemic [2] [3].

Taking into account that the main factors for developing overweight and obesity are related to healthy nutrition habits and physical activity, which are considered non-healthcare determinants of health, the problem for detecting risks of developing the condition or for following-up a patient that is under treatment is that the information containing the main risk factors does not often appear in an Electronic Health Record (EHR). Therefore, complementing the EHR information with this kind of information should be one of the main focuses for improving the evolution of this condition. The HHRs to be developed by the CrowdHEALTH project may be very useful to cover up this issue.

In addition, in the Spanish adult population (25-60 years) the rate of obesity is 14.5% while overweight is 38.5%. That is, one in every two adults has a higher weight than recommended. Obesity is more frequent in women (15.7%) than in men (13.4%). It has also been observed that the prevalence of obesity increases as people's age increases, reaching 21.6% and 33.9% in men and women over 55 years of age, respectively. Among the main causes are the greater consumption of hypercaloric foods (with high fat and sugar content) and less physical activity.

Among the goals of the current public health policies there are mainly the promotion of good nutrition habits and physical activity and the promotion of systematic detection of overweight and obesity, and specially the early detection of obesity to contribute to the reduction of

associated morbidity. However, public health policies do not appear to have had the expected effect.

Currently, overweight and obesity diagnosis are commonly underestimated [4] [5] [6] [7]. For instance, at the Health Department of La Fe in Valencia, the number of patients with an overweight or obesity diagnosis and currently are alive is 6,830, which represents approximately 2% of the capita of the Hospital (current capita is 298,975 citizens). This number shows how this condition is underdiagnosed if we consider that the Public Health Statistics state that around 50% of the population is overweight or obese. This fact could be reinforcing that, even if obesity factors are well known the implemented public health policies, seem to be ineffective. Thus, we have to understand which data are crucial to drive effective policies with a possible focus on risk stratification of obese and overweight patients that is to be carried out in the framework of the CrowdHEALTH project.

3.1.2. Key Players/Actors

The main actors involved in this Use Case are:

- Endocrinologists, the endocrinology service is the main supporter of this action as they need the information feedback to improve their health care attention on obese citizens and patients;
- Data analysts and researchers, might benefit from the data as they will provide answers on different research questions regarding the problem of overweight and obesity;
- Public Health Institutions, data might provide evidence regarding public health policy making on the promotion of physical activity, nutrition habits and systematic detection of obesity and overweight;
- Primary care professionals, who are the first health care professionals to diagnose and treat obesity; and hospital managers, who require information and indicators to take evidence-based decisions.

The available data from the Information Systems of Hospital La Fe are divided into different datamarts that include the following domains: patient information, hospitalization episodes, emergency room episodes, hospital at home episodes, and morbidity. Additionally, there is partial information that can be used for outpatient consultations, laboratory results, and costs. Hospital La Fe is carrying out an effort to integrate data from primary care and it will be partly available by January 2018. This data will include new clinical and anthropometric variables. The primary care data will provide also the ability to identify more patients diagnosed with obesity.

For the moment, the number of overweight or obese patients with a complete EHR and are currently alive is 6,830, which represent around 2% of the capita of the Hospital (current capita is 298,975 citizens). This is clearly an underdiagnosed condition since the public health analysis states that around 50% of the population has overweight or obesity. These patients

have been identified using their ICD-9-CM codes for obesity and overweight (278.00, 278.01, 278.02, 278.03).

The 6,830 identified patients have a total of 10,836 hospitalization episodes, 43,202 emergency room visits, 2,254 hospital-at-home episodes, 742,434 laboratory tests results, 553,832 diagnostics, and 140,893 outpatient consultations.

The data from the Hospital will be anonymized and de-identified upon ethical approval from the Ethical Board of the Health Research Institute it will be stored in a server at HULAFE.

3.1.3. Relevant Policies

The public health policies from the Public Health Institution are aligned with the Spanish Strategy for nutrition, physical activity and prevention of obesity (NAOS) from the Spanish Ministry of Health, Social Services and Equality.

The main objectives of this public heath strategy are (i) to promote policies and action plans aimed at improving dietary habits and increasing physical activity in the population; (ii) to raise awareness and inform the population of the positive impact of a balanced diet for their health and regular practice of physical activity; (iii) to promote nutrition education in family, school and community settings; (iv) to stimulate the practice of regular physical activity in the population, with special emphasis on schoolchildren; (v) to promote a framework of collaboration with food companies to promote the production and distribution of products that contribute to a healthier and more balanced diet; (vi) to sensitize the professionals of the National Health System to promote the systematic detection of obesity and overweight in the population; (vii) to monitor the proposed measures and the evaluation of the results obtained through the Strategy.

The available data from HULAFE and the Public Health Institutions are aimed at contributing to some of these objectives of the NAOS strategy. Mainly, in the use of data-driven models to promote a systematic detection of obesity and overweight in the population and in the clinical environment; but also at helping in the monitoring of the proposed measures and evaluation through the use of Key Performance Indicators. We also plan to use the results of some other use cases, mainly the SLOfit and the Myocardial infarction Use Case, of the CrowdHEALTH project to help in the promotion of good nutrition habits and physical activity. For instance, health professionals have raised concern in their ability to prescribe health activities, since they are not experts in the field, and they were asking for more formation in evidence-based promotion of physical activities.

3.1.4. Key Performance Indicators

There are up to 60 Key Performance Indicators (KPI:s) related to the aforementioned NAOS policies. The KPI appeared at the report Evaluation and monitoring of the NAOS Strategy: minimum set of indicators. From these KPIs we detect a lack of KPIs related to goal (vi): to

sensitize the professionals of the National Health System to promote the systematic detection of obesity and overweight in the population.

For this goal there are these related KPIs:

1. Overweight prevalence in adults: as the percentage of overweight people among the people surveyed, measured on adults (≥18 years old). Overweight is defined as people with a BMI ≥ 25 kg/m^2 and BMI < 30 kg/m^2.

2. Overweight prevalence in children: as the percentage of overweight children among the children surveyed (<18 years old). Overweight in children is defined as individuals with a Z-score > 1 standard deviation and ≤ 2 standard deviations in the Z-score tables of the WHO.

3. Obesity prevalence in adults: as the percentage of obese individuals among the people surveyed, measured on adults (≥18 years old). Overweight is defined as people with a BMI ≥ 30 kg/m2.

4. Obesity prevalence in children: as the percentage of obese children among the children surveyed (<18 years old). Obesity in children is defined as people with a Z-score > 2 standard deviations in the Z-score tables of the WHO.

5. Percentage of healthcare professionals that have received accredited training on healthy nutrition: as the number of healthcare professionals participating in accredited training among the number of healthcare professionals.

6. Percentage of healthcare professionals that have received accredited training on physical activity: as the number of healthcare professionals participating in accredited training among the number of healthcare professionals.

We, however, believe that a new KPI should be considered that takes into account the number of accurately identified patients with obesity and overweight among the true prevalence of obesity and overweight in the population.

This new KPI can be defined as:

7. “Increase of a 25% in the accurate identification of patients with overweight and obesity”.

We also expect to have a new risk stratification model for identifying the risk of overweight people to become obese people to alert and trigger the promotion of better habits. In addition, we also expect to have a model for identifying the remission of obesity to stimulate the ongoing acquisition of better healthy habits:

8. “Increasing the identification of overweight people at risk of obesity”.

9. “Decreasing the loss of follow-up of overweight and obesity treatment programs”.

As a technological KPI, we expect to deploy the systems for identifying overweight and obese patients and the risk stratification models at the pre-production stage of Hospital La Fe. This new KPI can be defined as:

10. “Deployment of the data-driven technologies at the pre-production stage of Hospital La Fe”.

Given that a more in-depth analysis of the data is needed, we believe there is scope in the data to improve the systematic detection of obesity and overweight. This would also allow covering the first two KPIs –we are not dealing with children data- by aiding in their computation.

To summarize, this Use Case intends to give help in the KPIs 1, 2, 7, 8, 9, and 10.

3.1.5. Expected Outcomes

The expected outcomes from CrowdHEALTH in this use case are related to:

1. Improved management and detection of obesity. This includes the systematic detection of people with obesity and overweight and the detection of bad nutrition and activity habits to promote on these citizens better habits.
2. Detection of groups of citizens with greater propensity for obesity to guide public health policies.
3. Broaden the knowledge of health professionals through a catalog of physical activity resources and professionals in order to improve the prescribing of physical activity. This should be combined with the knowledge, experience, and findings of other partners of the consortium.
4. Integrating new technologies for the monitoring of obese and overweight patients.

We expect to visualize the main KPIs, prevalence of obesity and overweight by groups of diets, individuals, sex, age and compare this to the prevalence of the National statistical surveys. An interesting visualization could be to visualize the prevalence by locations or regions.

3.1.6. Data Flow

The information of La Fe Health Department has two main sources (see Figure 1): the primary care flow, that uses the SIA (ABUCASIS) system and their own databases, and the secondary and tertiary care, that uses the ORION system with their own databases and the IRIS legacy system. The integration of both databases is not directly allowed for political reasons and the information for data analysis is not taken directly from the aforementioned databases. The common way a researcher or a data analyst can have access to the data for research goals is through a data warehouse that is designed and ready for favoring the analysis and efficient

presentation of processed data as information. This data warehouse keeps the data as SQL databases for on-line analytical processing and allows querying them and extract the relevant information as excel, ASCII text or CSV files that are perfectly fit to be treated with mathematical tools such as Matlab, R, or Python. The integration of the data from Primary Care within this data warehouse is currently being deployed but it is expected to be integrated by January of 2018.

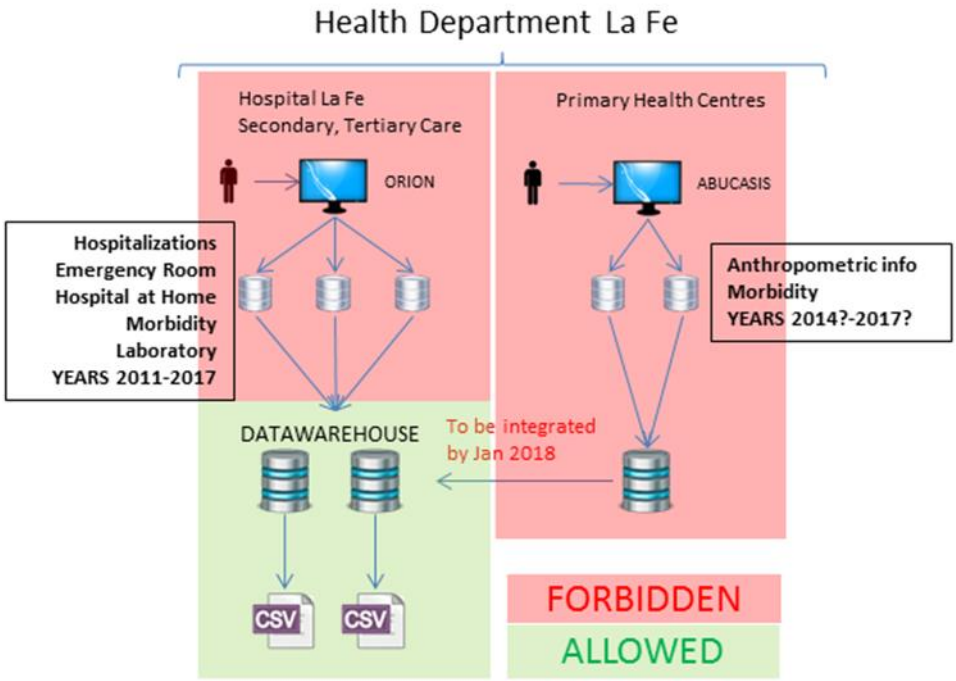

Figure 1: Data Flow HULAFE Use Case

3.1.7. Contributions to HHRs

Within HULAFE use case scenario, several contributions to HHR will be provided.

1. HHR Clinical Attributes
   a) Laboratory test results: glycosylated hemoglobin, Microalbumin/creatinine ratio, Glucose, Blood Urea, Creatinine, Albumine, Calcium, Sodium, Potassium, Transferrine, Troponin T, arterial CO2, arterial O2, hemoglobin, hematocrite, venous CO2, venous O2, Pro-BNP, Ferritine, Transferrine saturation Index, C-Reactive protein, Forced Expiratory Volume, arterial Ph, venous Ph, Total cholesterol, Low density cholesterol LDL, High density cholesterol, GOT transaminases, GPT transaminases, TSH thyroid, Free T3, Total T3, Free T4, Total T4.
   b) Hospital resource utilisation: outpatient consultation, emergency room visits, hospitalizations, hospital at home episodes.

c) Morbidities

2. HHR Nutrition

If available, there will be data of nutrition habits from a randomized sample of the children population with 600 individuals

3. HHR Life Style And Social Aspects

Socio-demographic data may be available: age clusters, sex, region of residence.

4. Relationships Among Citizens or HHRs

3.2. BioAssist – Chronic Disease Management

3.2.1. Brief Introduction

The BioAssist use case will capitalize on the company's ongoing piloting activities throughout Greece. Patients enrolled in BioAssist's pilots are provided with a tablet, with the BioAssist application pre-installed, as well as Bluetooth-enabled medical devices for measuring biosignals relevant to their condition (pulse oximeter, blood pressure meter, glucometer, spirometer, weighing scale, physical activity tracker). Each patient is supervised by a doctor enrolled in the platform and the doctor is capable of creating a personalized programme for measurements and medication intake for the patient. The patient performs biosignal measurements on a daily basis, following the doctor's instructions. Measurements are automatically transmitted to the BioAssist application and stored in the user's EHR, which also includes information such as medical test results, medications and allergies, and is accessible to their attending doctor, enabling them to make well-informed decisions and offer personalized care to their patients. Compliance to the programme is enforced via reminders. The system processes the data related to the schedule in real time and whenever a measurement exceeds a threshold that has been set by the attending doctor, the doctor is informed via a preselected communication channel (push notification, email, SMS etc.) Meanwhile, they can receive weekly automatically generated reports on their patients' status. The attending doctor communicates on a weekly basis with their patient, using the platform's videoconferencing functionality. The patient can also communicate through the system with their relatives and friends, while contact is encouraged with social networking features, such as photo and video sharing. All patients are supported by a 24/7 helpdesk, which provides timely response in case of emergency and is accessible to them with the touch of a single button.

The data collected within the BioAssist use case will enrich CrowdHEALTH's data sources with multiple and diverse types of information, covering both the clinical, the lifestyle and the social aspects of a patient's status, thus contributing instrumentally in the implementation of the HHR concept. The main objective of the chronic disease management scenario within the

project is to extend the existing monitoring solution with the CrowdHEALTH Platform's capabilities, will with the aim of creating a tool that will be available both to public health authorities and which will be able to provide meaningful insights into the collected data and empower potential policy makers in measuring the impact of relevant policies. At the same time, CrowdHEALTH's outcomes will also be highly useful to clinicians caring for groups of chronic patients, as is the case in the BioAssist pilots. Causal analysis and clinical pathway mining will allow for more effective decision-making with respect to the application of protocols, while patient profiling and forecasting will provide the means for better risk assessment and appropriate selection of treatment plans. In addition, patient clustering can offer insights from a macroscopic point of view, which may prove to be valuable in enhancing personalized care. Although there is already an operational pilot setup, which mainly involves patients with respiratory conditions, it is possible to create new deployments of the BioAssist platform within the piloting environments of other partners, either to equip their use cases with BioAssist's monitoring technologies or to collect additional data for a greater variety of chronic conditions, thus expanding the spectrum of policies that could be assessed within the use case.

3.2.2. Key Players/Actors

Role	Description
Patients	The patients enrolled in BioAssist's pilots are located in various regions in Greece and are mostly IPF and COPD patients, while some of them suffer from additional chronic conditions, including diabetes and CVD. They are the main providers of data for this use case, performing daily biosignal measurements. For some patients, lab test results may be acquired via web services from primary healthcare providers.
Clinicians	The clinicians involved in BioAssist's pilots are the attending doctors of the enrolled patients. They also provide data for the use case, by inputting information in their patient's EHRs (e.g. medications and lab test results). They are able to consult the collected data and receive weekly reports by the system to monitor their patients' health status.
Policy Makers	Entities that may create and apply health policies related to chronic disease management include public health authorities and insurance institutions. At the moment, there is no policy maker involved in the use case. However, clinicians in this use case monitor larger groups of patients and can therefore also be considered policy makers, in the sense that they apply clinical protocols, potentially deriving from relevant policies.

Table 1 Table of BioAssist Key Players/Actors

3.2.3. Relevant Policies

Our use case entails monitoring of patients with chronic conditions. Chronic diseases pose a considerable burden for national healthcare systems, as they increase costs in a multitude of ways, including health care use, medication, costly interventions and adding various indirect costs. The main goal in chronic condition care is not to cure, but to bolster functional status and quality of life, which requires a shift from traditional healthcare delivery models, away from fragmented service delivery and towards integrated care. Lately, many countries have attempted to create policy frameworks for chronic disease management, mostly aiming for better care service coordination and integration. The European Observatory on Health Systems and Policies has published a report which provides an overview of such interventions in 12 European countries and an assessment on the current status of chronic disease management in European health systems.[http://www.euro.who.int/__data/assets/pdf_file/0009/270729/Assessing-chronic-disease-management-in-European-health-systems.pdf] Examples of applied policies range from provision of informational material and patient coaching to the development of treatment pathways, decision support tools and employment of clinical information systems (e.g. electronic medical records and booking).

3.2.4. Key Performance Indicators

The scope of policy-making for chronic disease management is quite broad and not precise at the moment, therefore it is not easy to identify a definitive set of KPIs. Some common generic relevant KPIs are listed in the following table.

KPI	Description
Number of visits to doctor/specialist	This indicator relates to pressure on the health care system. While public health authorities usually keep records of patient visits, these records may not always be complete (e.g. visits to private physicians). This information could be derived for individual patients from the data collected within the use case, as doctors are able to schedule patient visits on the BioAssist platform.
Number of prescribed medications	This indicator relates both to healthcare costs, as well as to treatment results. This information can be derived from the data collected within the use case.
Number of (re)admissions to the hospital	This indicator relates both to healthcare costs, efficiency and the patients' actual health status.
Number of emergency room visits	This indicator relates both to healthcare costs, efficiency and the patients' actual health status.
Number of displacement by ambulance	This indicator relates both to healthcare costs, efficiency and the patients' actual health status.

KPI	Description
Level of patient adherence	This indicator refers to patient compliance with respect to treatment and self-management. There are numerous ways to estimate and interpret patient adherence. Considering the data that is available within the BioAssist use case, patient adherence for an individual may be calculated as the percentage of reminders (for medication intake, biosignal measurements, etc.) that the patient has responded to, at the appropriate time. For a population, this may be calculated as the median of the individual scores.
Life expectancy	There are many ways to estimate this indicator and difference approaches may be applied for different chronic conditions.
Grade of patient satisfaction	While increased patient satisfaction is one of the aims of policy making and there are various ways to calculate this indicator, the data collected within the use case cannot be used to extract such information.

Table 2: BioAssist KPIs

As mentioned in the table, some of these KPIs may already be available to policy makers (e.g. through public hospital records). Besides commonly used indicators such as the aforementioned, there is other information that can be extracted from the BioAssist use case and used for policy impact assessment with regard to health care results and quality of life.

Indicator	Description
Clinical Status	Combined analysis of the data collected within the use case can provide insights into patients' health status, allowing for assessment of policies that aim for improved health outcomes. Depending on the chronic disease, this may translate to specific indexes, such as the BODE Index for COPD.
Patient Behavior	Analyses of data on compliance, physical activity, social engagement and mood, can highlight outcomes of policies that aim to influence patient behavior and encourage a healthier lifestyle and better self-management.

Table 3: Other BioAssist KPIs

These indicators shall be defined in more detail in the course of the project.

The proposed indicators can be extracted from BioAssist use case. All data in our datasets can be easily visualized (e.g. in graph or table format) e.g. per type of biosignal. However, specific data visualization requirements shall become clear when the aforementioned indicators are properly defined. In general, visualization of HHR clustering results and highlighting patterns in data and trends in measurements shall be useful. Moreover, the differentiation of different clusters with similar characteristics utilizing a colour code shall

constitute a useful tool to directly evaluate the trends among groups and an optimal means to highlight immediately the impact of the various policies.

3.2.5. Expected Outcomes

We expect that CrowdHEALTH will bestow added value to patient monitoring technologies, transforming these into tools that support evaluation assessment with regard to attributes of a population that are currently difficult to examine, and providing a link between public health authorities and patients. The aim of the technologies already applied within our use case is to enhance patients with chronic diseases quality of life, encourage proactive care, and to offer efficient support in potentially dangerous situations. Furthermore, continuous monitoring of patients' vital signs and physical activity levels can supply doctors and other stakeholders of the healthcare ecosystem (e.g. providers, insurance institutions, etc.) with actionable data, allowing for enhanced decision-making and early detection of risks or deterioration of an individual's health status. Extending this scenario by exploiting the data analysis capabilities provided by CrowdHEALTH on collected data has the potential to equip policy makers (public health authorities and insurance institutions) with a tool that allows for measuring the impact of relevant policies, in terms of actual results on a population's health and quality of life. Examples of this could be training programmes for patient self-management support or behavioral modelling, incentives to motivate medication adherence, health insurance policies offering frequent medical examinations, investments in the promotion of active living and large-scale adoption of technologies that support independent living.

Analysis of the provided data may offer insights on a variety of questions, such as:

- Is an individual's/population's health status and quality of life improving/deteriorating?
- Is there correlation between user engagement with the provided technologies and their health status/quality of life?
- Is there correlation between social engagement and health status?
- Are there common attributes between individuals with improve/worsened health?

In addition, the platform's causal analysis and forecasting features will enable doctors to more effectively monitor disease progression and the evolution of a patient's medical condition, as well as the effects of interventions on individuals or groups of patients. Exploiting CrowdHEALTH's capabilities for patient profiling, clinicians will be able to perform risk assessment for their patients and accordingly form or alter treatment plans. CrowdHEALTH's results will be highly beneficial to the BioAssist platform, by providing tools that allow for extensive and holistic assessment of patients' condition and dynamic adaptation of clinical pathways.

3.2.6. Data Flow

There are approximately 100 patients enrolled in BioAssist's pilots. All patients perform daily measurements of their biosignals at home, according to their attending doctors' instructions. Each patient records at least two types of biosignals. The measurements are automatically captured by the BioAssist application and stored in the patients' EHRs on the BioAssist Cloud Platform. Attending doctors can view their patients' EHRs through a web interface and can add data, such as notes, medications, allergies, etc.

3.2.7. Contributions to HHRs

1. HHR Clinical Attributes

The BioAssist use case collects data on the following attributes:

- Biosignals (heart rate, SpO2, blood pressure, blood glucose, spirometry, weight)
- Physical activity
- Medication
- Lab test results
- Allergies
- Conditions (ICD-10)
- Demographics (age, gender, height)
- Self-reported Mood the patients indicate how they feel everyday by touching the respective icon on the screen– the options are worse, bad, good, better)
- Compliance (based on reminders)

2. HHR Sensors

The BioAssist platform is compatible with the following devices, which are used by patients enrolled in the pilots.

Device Type	Model		Measurement
Pulse oximeters	Jumper JPD-500F iHealth Air	 	Heart Rate (bpm) Oxygen Saturation (%)
Blood pressure monitors	Beurer BM85 iHealth Track	 	Heart Rate (bpm) Blood Pressure (mmHg)
Weighing scale	iHealth Core		Weight (kg)
Spirometer	MIR Spirobank Smart		FEV1 (L) Peak Flow (L/min)
Glucometers	iHealth Gluco Ascensia Contour NextOne	 	Blood Glucose Levels (mg/dL)
Activity tracker	H10pro		Step Count

Table 4: BioAssist Devices

3. HHR Stress

BioAssist does not directly collect data about stress. However, certain biosignals (blood pressure, heart rate) can provide the means to measuring the stress level of an individual.

4. HHR ADL

BioAssist does not directly collect data about ADL. However, daily activity levels and social interaction can be considered useful information regarding ADL.

5. HHR Life Style and Social Aspects

Within the BioAssist use case, the following information is collected:

- Mood (self-reported) the patients indicate how they feel everyday by touching the respective icon on the screen– the options are worse, bad, good, better)
- Social interaction (Videocalls, number of people in contact list, photos/videos shared with contacts)

6. Relationships Among Citizens or HHRs

Based on the available data, similarities among citizens that may be of interest, and therefore should be represented, are:

- Location
- Conditions (based on ICD-10 codes)
- Prescribed medications
- Demographics (gender, age, etc.)

Identifying similarities among citizens with respect to matters such as adherence to medication and measurement schedule, engagement with the system, activity levels and social engagement can provide a base for clusterisation and profiling of certain groups of users according to behaviour.

- Physical activity levels
- Response to reminders (delay, number of missed reminders)
- System usage analytics (e.g. number of sessions/week, frequency of measurements, frequency and duration of video calls, number of photos/videos uploaded

3.3. CareAcross-Coaching for cancer patients

3.3.1. Brief Introduction

The experience of being a cancer patient is a very challenging one. From the moment of diagnosis, patients' lives changes dramatically: the tumour itself brings about many negative consequences which affect their quality of life. Furthermore, demanding treatments, or combinations thereof, pose a significant burden on their overall health. Thankfully, research efforts and improvements in the treatment efficacy have led to higher survivorship rates and longer survival duration. On the other hand, these come at a substantial cost: worse quality of life.

At the same time, patients are increasingly becoming more active in their care. In particular, they are not satisfied with just undergoing surgery, receiving radiotherapy or chemotherapy, and taking more pills: they want to retain some level of control over their daily living. This is frequently expressed through their desire to improve their lifestyles. For example, many newly diagnosed patients make drastic changes in aspects of their lives including nutrition, exercise, social habits, etc.

The above parallel transitions are broadly understood, but not adequately tracked or documented. More specifically, although patients' quality of life is often reflected through their side-effects, it is not clear how well they are prepared for them, and how they report them. In addition to this, it is not clear how ready they are to make lasting improvements in their lifestyles.

This situation is clearly in dire need of data, but also of data analysis.

The Use Case is driven by the data entered by cancer patients on the secure CareAcross web platform. More specifically, patients provide information on their diagnosis, treatment, co-morbidities, health behaviours and side-effects. The platform then provides medical information based on this input, and attempts to "coach" the patients into behaviour change based on medical research.

An important and differentiating aspect of these datasets is that they are holistic in nature: they comprise of a variety of clinical (diagnosis, treatments) and non-clinical (health behaviours, side-effects) inputs.

Through the CrowdHEALTH platform, such data can be analysed in order to identify potential causal relationships between specific data points. For example, such a relationship may be that of a specific treatment and decreased adherence to coaching. Furthermore, it may enable predictions of future behaviours; for instance, that a patient with specific diagnosis will be less likely to report a specific side-effect.

Such analyses are very important for patients, for healthcare professionals, but also for public policy makers. This is because, the nature of oncology and cancer care services is mostly confined to the clinic; on the other hand, patients have increased and prolonged support needs. This means that, while there are no specific policies established for the provision of medical information and online coaching, such an approach may be quite helpful. This is not restricted only to the benefit of individual patients; it may also be fruitful towards the improvement of resource allocation in the healthcare system, reduction of costs, and decrease in the emergency or otherwise "reactive" care provision. Therefore, the main policy making goals of this Use Case are:

- to determine whether online coaching is a possible option for cancer patients
- to identify any relationship between the receipt of medical information on side-effects, and the reporting of these side-effects.

Potential collaboration with another project partner

Moreover, the CareAcross Use Case aims to incorporate aspects of another project partner (DFKI, standing for Deutsche Forschungszentrum für Künstliche Intelligenz - the German Research Centre for Artificial Intelligence) for a subset of the patients involved in the CareAcross Use Case. This will be through the use of activity trackers and the corresponding interface for collecting such data. This collaboration will enable further enrichment of the collected datasets, through a separate source and modality of data.

3.3.2. Key Players/Actors

The overall key players/actors are the following:

- Cancer patients
- Policy makers (indirectly)

The cancer patients themselves provide the data. This data is provided in the form of structured questionnaires through a private and secure web platform, offering specific input options for patients. Patients fill in these questionnaires online, and are incentivised to do so because they receive guidance and scientific information based on the exact input they provide. It is important to note that while every effort is made for the patient population to be uniform, there is a possible inherent bias due to the level of access to the internet that the responders would have, as well as due to their ability and interest in engaging with such services online.

As part and outcome of the CRA Use Case, the policy makers would focus on patient-centred communications (including, among others, medical information, behaviour change programmes etc.). (Policy makers will not be involved in the use case per se.)

More specifically, policy makers would benefit from the data analysis performed by the CrowdHEALTH platform which would indicate whether online coaching is a viable option in terms of its adherence by patients. In such a case, they would use the initial results from this Use Case to identify the key determinants of this adherence (through causal analysis as well as forecasting) in order to help compose suggested policies around online coaching for cancer patients. Similarly, they would benefit from the data points that the CrowdHEALTH platform would analyse around the impact of the provision of medical information on the quality of life of the patient (as exhibited through the reporting of side-effects). Such analysis would enable policy makers to recommend more specific guidelines on medical information "prescription" when it comes to the side-effects of cancer and its treatments.

These policies, while initially focused on the cancer patient population, may be extended to other patient populations, perhaps through additional research.

3.3.3. Relevant Policies

There are no current explicit policies on "medical information" (i.e. healthcare related information concerning specific diseases that is based on medical/clinical findings)" or behaviour change/coaching programmes. There are some limited "patient information" best practices in some locations (ranging from the clinic level to the national level, e.g. Patient Information Forum in the UK https://www.pifonline.org.uk/). However, they are very disparate and rarely are they aligned.

3.3.4. Key Performance Indicators

In order to analyse the data and evaluate their impact on the potential policies being discussed, the following KPIs will be calculated:

1. % of users who adhere to the coaching advice: This will reflect the extent to which patients actually adhere to the advice they receive from an online medium. This will be calculated as a percentage of patients who, upon receipt of specific advice, follow that advice. The latter will be based on the self-reported data subsequently provided by the patient. For the patients who also use an activity tracker, this combined part of the Use Case will allow us to conclusions on the impact of a physical wearable device on adherence.
2. % of users who remain engaged with the platform: This will reflect the extent to which patients actually use the platform for coaching. In other words, and given the nature of public policies, it is the long-term engagement with such a medium that is of importance. This becomes even more critical given the absence of a professional mediator or a formal trigger from the healthcare system. For the patients who also use an activity tracker, this combined part of the Use Case will allow us to draw conclusions on the impact of a physical wearable device on engagement.
3. % of users who report specific side-effects, in case they have received relevant education beforehand: This will reflect the extent to which the receipt of relevant

medical information affects patients' reporting of side-effects. Such a finding, when compared to the average (expected) reporting frequency of each side-effect, will allow us to draw conclusions on the side-effect over-reporting as well. Further causal analysis and forecasting will enhance our understanding of the impact of relevant education receipt on different populations of the same condition (e.g. different diagnoses, different treatments, etc). For the patients who also use an activity tracker, this combined part of the Use Case will allow us to draw conclusions on how differently those with a physical wearable device perceive medical information/education.

4. % of users who report specific side-effects, in case they have not received relevant education beforehand: This will reflect the extent to which patients who have not received medical information about side-effects actually report side-effects. Causal analysis and forecasting will then enable us to determine frequent side-effects per diagnosis and treatment combination, as well as any other combination (e.g. unhealthy daily habits).

3.3.5. Expected Outcomes

The main outcomes will be answers to the questions:

1. Does online coaching support behaviour change? This will be assessed via the KPIs (1), (2) above. Unfortunately, there is no specific "baseline" for this, so the assessment will be mostly qualitative.
2. Can online medical information help prevent side-effects? This will be assessed via the KPIs (3), (4) above.
    - These outcomes will be delivered through the following processes within CrowdHEALTH: Data analysis among the datasets in order to generate cohorts across (a) adherent patients (b) non-adherent patients (c) patients who receive medical information (d) patients who do not receive medical information (e) patients with a fitness device.
    - Causality analysis in order to determine the determinants of the above KPIs and, subsequently, their outcomes
    - Forecasting in order to enable posing fundamental questions for policy makers that would allow them to identify potential policies and even test them based on prior data.

3.3.6. Data Flow

The data is entered by the patient. The patient subsequently receives information and/or advice, and provides feedback about whether the advice has been followed (if applicable).

The data entered includes the following:

"Static" data:

- Diagnosis
- Treatment
- Co-morbidities

"Updated" data (repetitive):

- Nutrition intake around specific categories of food
- Side-effects.

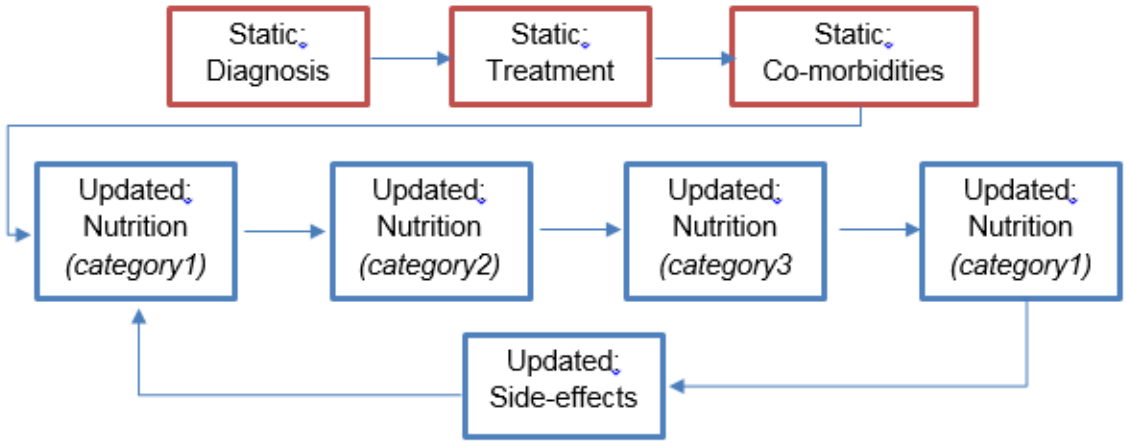

Figure 2: CRA Data Flow

3.3.7. Contributions to HHRs

1. HHR Clinical Attributes

Within CRA scenario there is no input from attributes coming from clinics or any other healthcare centres, meaning that there is not direct integration of data from EHRs and other resources. The clinical attributes such us diagnosis, treatment and co-morbidities is used in the scenario they are inserted in CareAcross platform from the patient.

2. HHR Nutrition

The CareAcross Use Case collects information about:

i. Red Meat
ii. Poultry
iii. Fish
iv. Dairy
v. Eggs

vi. Legumes
vii. Fruits
viii. Vegetables
ix. Potatoes
x. Bread

The source of this data is the patient; based on the flow above, this is repetitive and updated periodically.

The below table summarises the above inputs:

Category	Food item	Unit/week
1	Red Meat	Portions
1	Poultry	Portions
1	Fish	Portions
2	Dairy	One unit=1 glass of milk, 1 cup of yoghurt, or 2 slices of cheese
2	Eggs	Units
2	Legumes	Portions
3	Fruits	One unit=1 Fruit, or a handful of berries/grapes, or a glass of fresh juice
3	Vegetables	One unit=1 Cup
4	Potates	Units
4	Bread	One unit=1 slice

Table 5: CRA inputs about HHR Nutrition

3. HHR ADL

The CareAcross Use Case captures information about side-effects as related to cancer

- Fatigue
- Nausea
- Vomiting
- Constipation
- Mouth sores
- Dry mouth
- Changes in taste
- No appetite
- Pain in stomach
- Problems swallowing

3.4. University of Ljubljana-Physical fitness, physical activity and obesity

3.4.1. Brief Introduction

Our use case is aimed at providing direct support to school physicians and pediatricians to help them in early detection of children with increased health risks, linked to poor physical fitness and obesity, in their decision-making and in surveillance of the effects of any health intervention programs in which those children will participate. Our aim is to have teachers, parents, school physicians and pediatricians working together and setting up health intervention programs that could run simultaneously within primary health care institutions and schools according to joint strategies.

The current data flow between school physicians and pediatricians is very limited and their communication most often non-existent. Currently, schools and health institutions have no official channels of communication established, accept for the scheduling of health check-ups in local health institutions. The school physician also does not communicate with schools and teachers about specific children, who are diagnosed with any health-risks and only sends a statistical report on the check-ups back to school principals. When dealing with obesity, school physicians might only direct and invite the parents of morbidly obese children to consult any health specialist, while obese and overweight children are mostly being ignored. A school physician also has no knowledge about the somatic and physical fitness development of a child apart of weight and height gathered every second year. The same goes for the pediatrician as well which means that they lack a lot of information that could help them better assess health risks or understand their symptoms. With constant communication and data access of school physicians and children's pediatricians with parents and teachers, the utilization of data, gathered by annual SLOfit testing in schools, it would be possible to overcome the current problems of data isolation and boost the possibility of joint health-related interventions, managed by teachers and school physicians or by parents, pediatricians and teachers. In the initial phase, the school physicians and pediatricians would be able to access the SLOfit data of children only directly through SLOfit database but after the setup of the e-health records within the health sector, the data from the SLOfit database could be pulled directly into e-health system. A concrete example would be exercise and nutrition programs for morbidly obese children in primary health care institutions or school-based sports programs that enroll overweight and physically ineffective children. The system, developed within our use case would enable fast, sustainable and controlled flow of information between teachers, parents, children, school physicians and pediatricians.

The use case is going to be implemented as a pilot in one local environment of the Škofja Loka city in Slovenia. Škofja Loka is a typical administrative centre, Slovenia, in which we can include all school-going age groups. It has two secondary schools which enrol youth from 15 to 19 years of age from the city and the municipalities surrounding it, and three primary schools which enrol children from 6 to 15 from the city itself. We will include all three primary and two secondary schools (around 2000 students) for which we will set up a cohort database on physical fitness and somatic development for every child from age 6 to 19 and will

supplement this cohort data for every child with cross sectional data on physical activity, sedentariness, sleep, resting heart-rate, socio-economic status, and parental physical activity. We will also include several school physicians who are going to get access to the SLOfit data on children's somatic and physical fitness development with permission of the parents.

At the school level, all children in the pilot environment, school principals, teachers, and parents will be involved in the pilot. The main beneficiaries will be parents and children, who will get thorough and graphically supported information on their development, current status and future predictions of physical and motor development, and current status of all the other indicators noted above. At the same time, we will provide school physicians with access to the national SLOfit database, which includes annual data on somatic development and physical fitness development of the entire population of children in Slovenia, which will considerably improve their ability to assess health risks and their diagnostics at regular health check-ups or enable early detection of health risks, related to poor physical fitness or obesity, and include those children in appropriate intervention programmes.

The parents of the primary-school children and the secondary-school students themselves already provided positive consent for their data to be analysed and stored in the SLOfit database.

Within the CrowdHEALTH project, the use case will utilize its platform to produce the tools for data visualization that are going to be used as a part of the HHR. The use case is going to provide data on physical fitness and physical activity to supplement the data on nutritional status of children and enable the construction of obesity risk assessment and developmental prediction models of somatic and physical fitness development. In this way, HHR is going to be supplemented by a possibility to include one or more physical fitness or physical activity components into the assessment of children's future health risks. The CrowdHEALTH platform will provide the reporting system for various end-users ranging from individual users/patients to health policy makers/ministries and governments (see also section 3.4.5 Expected outcomes).

3.4.2. Key Players/Actors

The main players in the use case are the Faculty of Sport of the University of Ljubljana, primary and secondary schools, parents and children, teachers, school physicians and paediatricians, local health institutions, municipalities, National Institute of Public Health, Ministry of Health and Ministry of Education, Science and Sport (see Figure 3).

The SLOfit data is provided by the Faculty of Sport, University of Ljubljana who gathers the data from all Slovenian primary and secondary schools every April. This data is also stored at the Faculty of Sport. The SLOfit team at the Faculty of Sport manages the database, prepares reports for every school and every child, consults the data with schools and the Ministry of Education and prepares annual reports. Annual reports are publicly available every year by the end of November.

In Slovenia, every school has a designated school physician while every child has a designated paediatrician who is most often not the school physician. The role of the school physician is to regularly check up on the children (every 2 years in primary school, at the beginning of secondary school and at the end of the secondary school) and vaccinate them according to the national regulations. When a child progresses from primary to secondary school, the school physician is changed and the parents are required to have the health records sent from the primary-school physician to the secondary-school physician, which very often does not work out. Children's paediatricians act as their primary physicians. In the case of acute or chronic illness, children see their paediatricians which are responsible for their treatment or delegation of their case to other health specialists.

National Institute of Public Health acts as the coordinator of any activities planned by the Ministry of Health and is responsible for implementation of any health policies from the national level to the local settings. Its regional centres coordinate any activities on the local level.

3.4.3. Relevant Policies

Our case is directly linked with national and international policies, related to stopping childhood obesity and increasing physical activity: the WHO Global Action Plan for the Prevention and Control of NCDs 2013-2020 has a goal of 10% relative reduction in prevalence of insufficient physical activity, and to halt the rise in diabetes and obesity by 2020. On the national level a relevant policy to which our case adheres is described in the National Program on Nutrition and Physical Activity for Health 2015-2025.

3.4.4. Key Performance Indicators

Currently, one of the biggest problems for policy makers in the area of Public Health in Slovenia is their inability to set up evaluation or surveillance systems through which they could evaluate and monitor the effects of the implemented policies. When it comes to health prevention, most of the implemented policies are poorly informed by any hard data and rather inappropriately evaluated (for example, the success of the new policy on which the intervention programs for reducing obesity is evaluated by the number of participants and their liking of the program rather than with their progress in fat mass reduction). Our proposed KPIs are in the following table.

KPI	Description
Physical fitness index	An overall evaluation of physical effectiveness in percentiles according to age and gender.
Health-related fitness index	The estimation of fitness based on motor tests 600 m run, bent arm hang, and sit ups, which gives the estimation of health-related fitness in

KPI	Description
	percentiles according to age and gender.
Obesity index	Nutritional indicator, based on BMI and triceps skin fold, which gives the estimation of obesity in percentiles according to age and gender.
Prevalence of pre-obesity and obesity in children	Calculated as percentage of children between 6 and 19, scoring between 1 and 2 SD or over 2 SD above national BMI median.
Prevalence of children with low and very low physical fitness	Share of children, scoring between -1 and -2 SD or -2 SD below national physical fitness index median.
Stature	Measured in cm and in percentiles according to age and gender.
Weight	Measured in kg and in percentiles according to age and gender.
Triceps skinfold thickness	Measured in mm and in percentiles according to age and gender.
Arm plate tapping	Measured in repetitions in 20 seconds and in percentiles according to age and gender.
Standing broad jump	Measured in cm in percentiles according to age and gender.
Obstacle course backwards	Measured in seconds in percentiles according to age and gender.
Sit ups	Measured in repetitions in 1 min in percentiles according to age and gender.
Stand and reach	Measured in cm and in percentiles according to age and gender.
Bent arm hang	Measured in seconds in percentiles according to age and gender.
60 m dash	Measured in seconds in percentiles according to age and gender.
600 m run	Measured in seconds and in percentiles according to age and gender.
Physical activity index	Derived from physical activity and sedentariness, which gives the estimation of physical activity.

Table 6: University of Ljubljana KPIs

Proposed KPIs can be extracted from our case data with relative ease. Such data is also fairly easy visualized and can be used in “before and after” manner very effectively.

The data analysis will provide answers on the patterns of physical and motor development and their interrelatedness, and the relation between physical activity, sedentariness, SES, parental physical activity and physical fitness. The data analysis should provide answers on how much physical activity is necessary for retaining adequate level of physical fitness, and what ratios of physical activity and sedentariness still provide sufficient physical fitness.

3.4.5. Expected Outcomes

We expect the CrowdHEALTH to create a basis for implementation of policies that enable linking school and health data for early interventions monitoring and evaluation. We would like to visualize the individual growth trends, physical fitness development trends, nutritional development trends, physical activity trends, adult stature prediction, adult weight prediction, adult physical fitness prediction, adult obesity-related health risks prediction for the 2000 individuals, included in the use-case from year 2004 onwards. Afterwards, the entire population of children and youth in Slovenia will be included (around 240,000).

3.4.6. Data Flow

The SLOfit data (physical fitness) is gathered every April with field testing in all (about 600) primary and secondary schools in Slovenia. Schools enter the data in spreadsheet (Excel or dedicated data entry program) and currently send this data by e-mail to the Faculty of Sport, University of Ljubljana (FSUL) where it is checked, cleaned and analyzed (The complete data flow diagram for our use case is given in Figure 3. In 2018 (pilot) and 2019 data will be entered / imported directly in SLOfit database by school administrators. After import, indexes (BMI, overall fitness index) and percentiles (according to sex and age) of fitness data are computed within SLOfit web application. On basis of these analyses, FSUL will produce three types of reports: school reports that will be sent back to schools, individual reports that will be sent directly to children, their parents and to children's pediatricians (with parental consent), and national report with aggregated data sent to Ministry of Education and National Institute of Public Health. In the pilot, we are going to use also web-based physical activity questionnaires and wearable energy expenditure monitors to supplement the existing profiles of children also with physical activity and somatic status data. The summaries of school reports are going to be presented also to municipalities to inform them on the current physical status of children in individual schools and possibilities of improvement, linked to accessibility of green spaces, playgrounds, safe school walking and cycling routes and sports infrastructure.

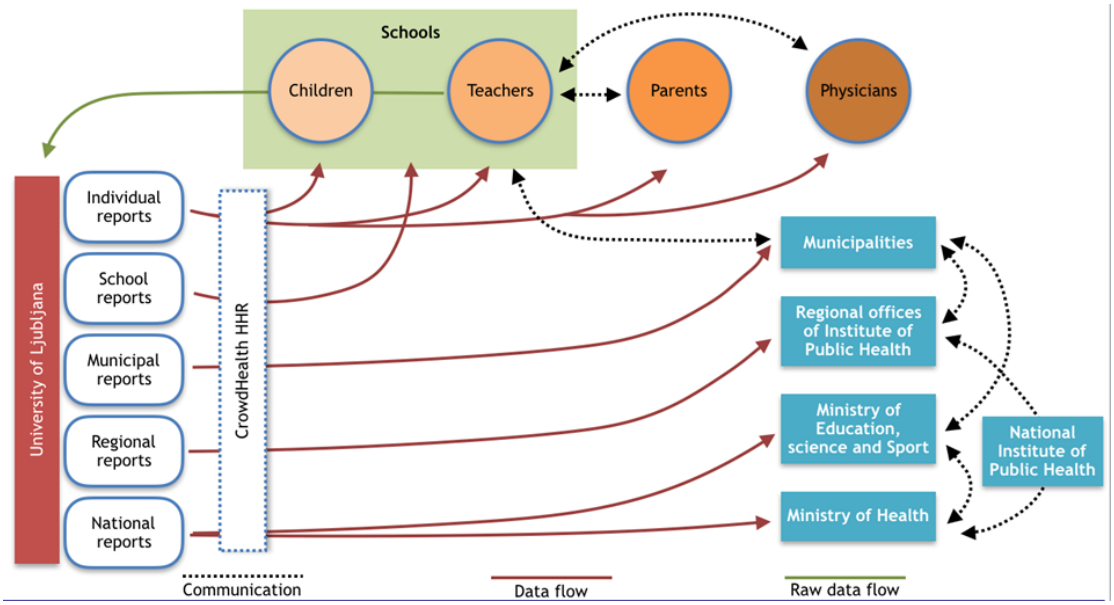

Figure 3: Data flow between the key players

If the National Institute of Public Health succeeds in establishing e-health record in which school physicians and pediatricians will be able to transfer child's health records such as blood lipids, blood pressure, etc., this information could be merged with SLOfit data in CrowdHEALTH.

3.4.7. Contributions to HHRs

1. HHR Clinical Attributes

   - Obesity
   - Physical effectiveness
   - Muscular fitness
   - Aerobic fitness
   - Physical activity

2. HHR Sensors

Volume of physical activity from physical activity trackers (measured by multi sensor devices on subsample of 200 students), monitored for longer period, 24 hours per day. The sensors will be measuring daily active energy expenditure and duration of physical activities of different intensities.

3. HHR Stress

Sleep quality (measured by multi sensor devices within pilot on subsample of 200 students).

4. HHR ADL
   - Volume of sedentariness (measured by questionnaire on sample of 2000 students and by multi sensor devices on subsample of 200 students).
   - Volume of physical activity (measured by questionnaire on sample of 2000 students and by multi sensor devices on subsample of 200 students).
   - Sleep duration (measured by questionnaire on sample of 2000 students and by multi sensor devices on subsample of 200 students).

5. HHR Lifestyle and Social Aspects
   - Socio-economic status (measured by questionnaire on sample of 2000 parents of students).
   - Parental physical activity (measured by questionnaire on sample of 2000 parents of students).
   - Academic achievement (measured by questionnaire on sample of 2000 students).

3.5. Karolinska Institutet

3.5.1. Brief Introduction

According to the National Swedish Board of statistics, it has been observed in the last two decades a dramatic reduction of mortality in acute myocardial infraction. In 1987, 39.000 individuals were affected by a heart attack of which 30.000 were hospitalized. In 2010, 32.000 individuals had diagnosed with myocardial infraction of which 26.000 were hospitalized. The number of deaths from myocardial infraction as the underlying or contributing cause has fallen from 18.000 individuals 1987 to 9.000 individuals.[Socialstyrelsen, Nationella riktlinjer för hjärtsjukvård, 2015a] Despite these results, cardiovascular diseases remain the leading cause of death in both men and women in Sweden. In 2010, cardiovascular disease was the underlying cause of death (40%) where 42% of them diagnosed as ischemic heart episodes. The health care costs in 2010 for cardiovascular diseases, according to the Economic Institute of Health Care in Lund amounted to 61.5 billion Swedish kronor of which health care costs totalled to 41 %, informal care for individuals 30% and 29% for production costs.[Socialstyrelsen, Nationella riktlinjer för hjärtsjukvård, 2015a] The diagnosis and treatment of acute and chronic diseases today, is mainly based on scientific studies and most of the treatments have clearly proven effects in terms of improved survival, decrease the risk of relapse and quality of life improvement. Recommendations for new treatments take several years of investigation in the Swedish National Board of Health Care Policies. Still, there are variations between different hospitals and health care organizations in the use of diagnostic tests, drug therapy, catheter-based and surgical interventions, which have implications for the public health and public health economics.

Through the CrowdHEALTH platform, we aim to develop an enhanced digital information environment which will include integrated multitude data sources providing information that has not yet reached its full potential for cardiovascular diseases. The goal is to introduce a

new paradigm of Holistic Health Records (HHRs) providing a decision support tool to public health authorities for policy creation and co-creation, through the exploitation of collective knowledge that emerges from multiple information sources and its combination with situational awareness artefacts. Within the project, a new paradigm of electronic health records will be introduced and tested aiming to understand if processing enhanced electronic health records could assist the evaluation, creation and implementation of public health policies for cardiovascular diseases.

For this purpose, during the first phase of the project we are going to collect information from two databases in Stockholm County. The first one is the VAL database which includes EHRs from all the registered hospitals in Stockholm. The second source will be the SwedeHeart quality registry which contains EHRs only for patients diagnosed with cardiovascular disease. In the second phase and after the creation of HHR, this data will be processed by the health analytics toolkit of CrowdHealth platform aiming to visualize advanced analytics information to the public health policy makers. The processing of this data will be carried from health analytics toolkit and it will give the opportunities to the policy makers to understand better the effectiveness of current policies and to enhance the policy creation processes.

3.5.2. Key Players/Actors

Karolinska Institutet (KI) is the first key player in the CrowdHEALTH project. It is responsible to manage the data collection, aggregation and communicate with the other stakeholders that are involved. KI is responsible for hosting the extracted data in a secure environment under Institute's technical infrastructure and ensure the accessibility of this data from the CrowdHEALTH platform. It is responsible also to communicate the outcomes to the research society nationally and globally.

SwedeHeart quality registry (http://www.ucr.uu.se/swedeheart/) is the second key player in the project. SwedeHeart Database contains data from all patients with acute myocardial infarction, secondary prevention data in those younger than 75 (RIKS-HIA and SEPHIA). Moreover it contains data from all patients undergoing angiography, percutaneous coronary intervention (SCAAR), heart surgery (heart surgery registry) and percutaneous valve intervention (Percutaneous valve intervention registry). Swedheart quality registry it's also a key player in CrowdHEALTH project in terms of policy making around cardiovascular diseases. Nationally, policy makers responsible for cardiovascular diseases are using the annual reports and the published research studies from SwedeHeart research groups. Based on the findings they conduct new studies in order to examine how they can use this information for:

1. Evaluating the effectiveness of a current policy
2. Exploring the impotence of a current policy
3. The creation of a new policy

They will be the main users of the platform in our scenario and they will participate in all the stages that are included in the project for policy creation, evaluation and implementation as well as evaluating the CrowdHEALTH platform for its effectiveness based on scientific findings which show that the platform improves the policy making procedures in Sweden for cardiovascular diseases.

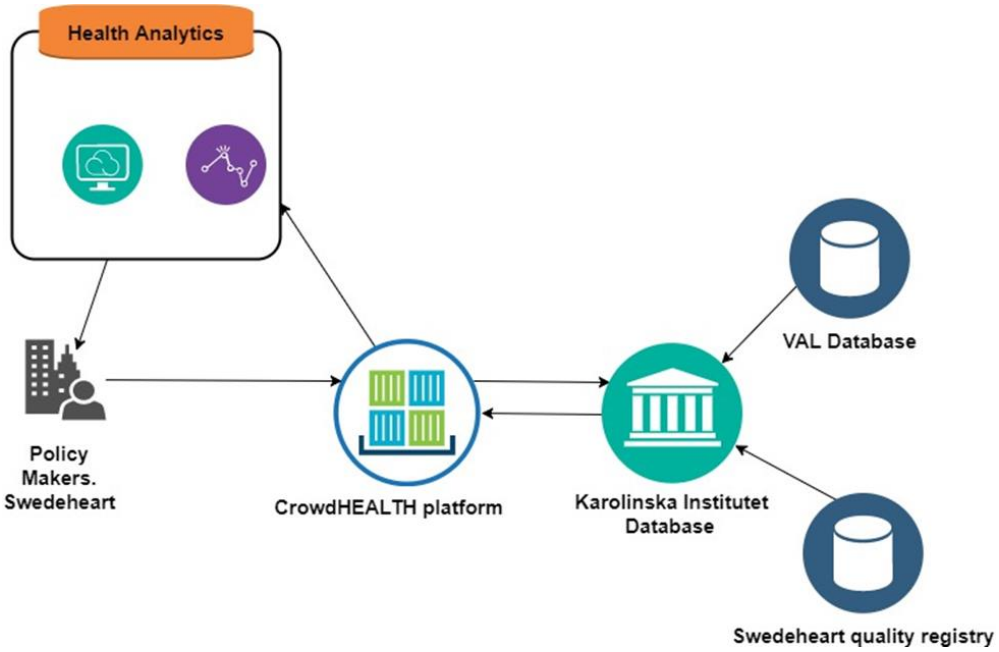

Figure 4: Key players/Actors flow diagram

3.5.3. Relevant Policies

Several policies will be examined based on the health analytics toolkit in terms of applicability of a new policy and the effectiveness of the current ones. The policies used will be based on:

1. Time scales (i.e. long- / short- term) policies.
2. Hospital based policies and differences for cardiovascular patients.
3. Population segmentation based on common disease diagnosis, age, gender.
4. Evolving risks such as stressful life style.

One example is to measure by the extent to which the findings allow for a better understanding of differences in treatments as well as factors that affect the treatment, and whether these lead or will to improve practices and guidelines around a targeted patients group.

3.5.4. Key Performance Indicators

Based on the most recent cardiovascular research studies, the KPIs are divided and will be examined within CrowdHEALTH project as follows:

Population based core indicators:

- Percentage of adult patients who have weight and height or waist circumference recorded
- Percentage of adult patients who have alcohol consumption recorded
- Percentage of adult patients who have smoking status recorded
- Percentage of patients who are current smokers and have smoking cessation counselling or a referral for counselling recorded
- Percentage of patients 40 years and older with no risk factors, or any adults with cardiovascular risk factors (e.g. hyperlipidaemia, hypertension, ischemic heart disease), who have had a fasting plasma glucose level recorded on the chart in the past three years.
- Percentage of patients older than 40 years of age (men) and older than 50 years of age (women) for whom a global risk assessment (e.g. Framingham model) has been recorded on the chart.

Hypertension Based Indicators

- Percentage of patients with an average systolic blood pressure of greater than 160 mmHg and/or a diastolic blood pressure greater than 100 mmHg, as determined on at least three separate visits, who have a diagnosis of hypertension recorded
- Percentage of adult patients whose blood pressure is 180/110 mmHg or greater, or 140/90 mmHg or greater and who have diabetes, chronic renal disease or target organ damage, who have a record on the chart of a second visit for blood pressure within two months of the first elevated blood pressure visit.
- Percentage of adult patients whose blood pressure is 180/110 mmHg or greater, or 140/90 mmHg or greater and who have diabetes, chronic renal disease or target organ damage on a second visit, who were labelled as hypertensive.
- Percentage of patients with an average systolic blood pressure of 160 mmHg or greater, or a diastolic blood pressure of 100 mmHg or greater with a recommendation for drug therapies recorded
- Percentage of patients with an average diastolic blood pressure of 90 mmHg or greater with a recommendation for drug therapies recorded on the chart if target organ damage is present or if they have independent cardiovascular risk factors (elevated systolic blood pressure, cigarette smoking, abnormal lipids, family history of premature cardiovascular disease, truncal obesity, sedentary lifestyle)
- Percentage of patient visits (for blood pressure follow-up) for those with hypertension whose blood pressure is above target (140/90 mmHg, or 130/80 mmHg for patients with diabetes or renal disease) with a plan of care for hypertension recorded on the chart that includes a change in dose or regimen of medications, and/or repeated education regarding lifestyle modification and/or planned reassessment.
- Percentage of patients identified as hypertensive, but who are at target blood pressure levels and who have had blood pressure recorded in the chart in the past six months.

Key performance indicators for chronic, stable ischemic heart disease

- The percentage of patients with ischemic heart disease who are taking acetylsalicylic acid or have a contraindication to, or side effects from, acetylsalicylic acid.
- The percentage of patients with ischemic heart disease who have had a myocardial infarction and are taking a beta-blocker or have a contraindication to, or side effects from, a beta-blocker.
- The percentage of patients with ischemic heart disease who are on an angiotensin-converting enzyme inhibitor, or have a contraindication to, or side effects from, an angiotensin-converting enzyme inhibitor.

Key performance indicators for congestive heart failure

- Percentage of patients with a diagnosis of congestive heart failure who have had an ejection fraction value recorded in the chart at least once.
- Percentage of patients with left ventricular systolic dysfunction (ejection fraction of less than 40%), whether symptomatic or asymptomatic, who are taking an angiotensin-converting enzyme inhibitor or an angiotensin receptor II blocker, or have a contraindication to, or side effects from, both an angiotensin-converting enzyme inhibitor and an angiotensin receptor II blocker.
- Percentage of patients with left ventricular systolic dysfunction - ejection fraction of less than 40% - who are taking a beta-blocker or have a contraindication to, or side effects from, beta-blockers.
- Percentage of patients with congestive heart failure on an angiotensin-converting enzyme inhibitor or an angiotensin receptor II blocker who have had potassium and creatinine levels recorded on the chart in the past year.
- Percentage of patient visits for congestive heart failure during which weight was recorded in the chart.

Key performance indicators for hyperlipidaemia

- Percentage of adult patients with one or more of the following who have lipid testing recorded on the chart every two years: diabetes mellitus; hypertension and/or risk factors, such as smoking or abdominal obesity and/or strong family history of premature ischemic heart disease; or evidence of symptomatic or asymptomatic coronary artery or vascular disease
- Percentage of patients with hyperlipidaemia for whom a therapeutic target, based on their global risk assessment and lipid profile, has been recorded on the chart.
- Percentage of patients with hyperlipidaemia who are at high risk for ischemic heart disease, for whom it has been recorded on the chart that pharmacological treatment was recommended immediately, concomitant with dietary and lifestyle changes.

3.5.5. Expected Outcomes

Based on the CrowdHEALTH technologies we expect to deliver a modern policy making framework that will facilitate policies evaluation and optimization through adaptive and incremental visualizations of simulations and outcomes of evidence based analysis of prevention strategies. The main expected outcome is to enhance the current processes of evaluating, creating and implementing public health policies for cardiovascular diseases in Sweden.

Moreover the policy making framework can provide added-value with respect to policy making through the identified clusters / segments of populations with similar characteristics and it will open new opportunities for personalized medicine, disease prevention, and effectively leading to a reduction in readmission rates for cardiovascular disease patients.

Finally, within CrowdHEALTH project a new modern schema of electronic health records is being proposed. During the project, this schema will be tested for its connectivity with other data sources (real time data or data from other research projects in Sweden) and explore the level of the records' enhancement and how this is helping on the scalability feature of CrowdHEALTH platform.

3.5.6. Data Flow

Karolinska Institutet will use two static data extracts from VAL database and SwedeHeart quality registry. The regional healthcare data warehouse GVR/VAL contains diagnoses (ICD-10), drugs (ATC), and other data related to consultations in primary and secondary care for more than 2 million inhabitants of the greater Stockholm area. The population selected for this project consists of all individuals in the GVR/VAL data warehouse that since 1997 have been diagnosed with cardiovascular disease. Swedeheart Database contains data from all patients with acute myocardial infarction, secondary prevention data in those younger than 75 (RIKS-HIA and SEPHIA). Moreover it contains data from all patients undergoing angiography, percutaneous coronary intervention (SCAAR), heart surgery (heart surgery registry) and percutaneous valve intervention (Percutaneous valve intervention registry).The population selected for this project consists of all individuals in the Swedeheart Quality Registry that since 1997 have been diagnosed with cardiovascular disease.

3.5.7. Contributions to HHRs

1. HHR Clinical Attributes

Based on the public health policy makers and clinicians several clinical attributes could be examined and integrated in to the holistic health record. At this stage we have listed the followings:

- Blood pressure
- Blood Sugar

- Body Weight
- Blood Lipids
- Prescription
- Other disease (Diabetes)

2. HHR Life Style and Social Aspects

Similar for the life style and social aspects the following information will be aggregated at this stage in the HHRs are the followings:

- Tobacco use
- Alcohol
- Stress levels
- Contraception Pills

3.6. DFKI-HULAFE and DFKI-CRA Integrated Use Cases

3.6.1. Brief Introduction

Initial, the DFKI Use Case is concerned with collecting nutritional data and activity data from citizens. This consists of having a low-threshold collection of what persons eat via the kitchen kit app, which includes different mechanisms in order to be as accurate as possible. First, when eating at home, it offers meal suggestions taking into account information about diets the persons need or want to follow, and supporting suggestions of alternative ingredients if necessary. In order to also collect the information when eating out, it will support to quickly gather what has been eaten there. With respect to activities, we will use available accelerometer and possibly blood pressure or pulse sensors (e.g. in smartwatches, Fitbit, etc.) to detect physical activities during the day. If possible, we may use additional sensor information to detect where a person spends time at home (e.g. by placing beacons in the room) or outside (e.g. from GPS). The gathered information will be the type and intensity of physical activities. As with the nutritional data, we will include in the kitchen kit app a part to complete the information on the physical activities (e.g., kind of activity, other non-detected activities) and social information.

The use case assumes that all participating persons have an initial holistic health record, where health information (especially those related to specific diets) as well as lifestyle, occupation, environmental and maybe geographical information is included. This information is collected in a questionnaire and may be updated during the participation in the pilot. Ideally each person will participate four times in a year in order to detect seasonal variations. The collected personal data and derived knowledge has foreseeably different degrees of reliability. This may result from known inaccuracy of used devices (e.g., pulse measurement) or just unavailable details, such as ingredients of meals when eating out. Hence, we will include a

mechanism to assess the accuracy of information and in the entries into the HHR include a confidence of the entries.

Since there is no possibility to get the electronic health records of any patient, not only because of data security issues but primarily because those do not exist for most regions and patients, the DFKI study will take place in cooperation with HULAFE and CRA. This way the electronic health record as well as the collected data are available for the patients.

In the integrated use case the patients under treatment at HULAFE will be selected for the study during the outpatient consultations and will be equipped with an activity and nutrition tracking devices developed and provided by DFKI. This allows complementing the health records kept at HULAFE with activity and nutritional data collected when they are outside the hospital to obtain holistic health records.

The participating patients shall be volunteers that will read and sign an informed consent following all the ethical recommendations and ideally they will participate as long as possible in the use case, but at least 12 months.

DFKI domain is concerned with collecting nutritional data and activity data from citizens. This consists of having a low-threshold collection of what persons eat via the kitchen kit app, which includes different mechanisms in order to be as accurate as possible. First, when eating at home, it offers meal suggestions taking into account information about diets the persons need or want to follow, and supporting suggestions of alternative ingredients if necessary. In order to also collect the information when eating out, it will support to quickly gather what has been eaten there. With respect to activities, we will use available accelerometer and possibly blood pressure or pulse sensors (e.g. in smartwatches, Fitbit, etc.) to detect physical activities during the day. If possible, we may use additional sensor information to detect where a person spends time at home (e.g. by placing beacons in the room) or outside (e.g. from GPS). The gathered information will be the type and intensity of physical activities. As with the nutritional data, we will include in the kitchen kit app a part to complete the information on the physical activities (e.g., kind of activity, other non-detected activities) and social information.

The use case assumes that all participating persons have an initial holistic health record, where health information (especially those related to specific diets) as well as lifestyle, occupation, environmental and maybe geographical information is included. This information is collected in a questionnaire and may be updated during the participation in the pilot. Ideally each person will participate four times in a year in order to detect seasonal variations. The collected personal data and derived knowledge has foreseeably different degrees of reliability. This may result from known inaccuracy of used devices (e.g., pulse measurement) or just unavailable details, such as ingredients of meals when eating out. Hence, we will include a mechanism to assess the accuracy of information and in the entries into the HHR include a confidence of the entries.

Let us now elaborate on the technical realization of this integrated use case.

HULAFE will have a local installation of the CrowdHEALTH platform at their site and with all necessary security protections. No access to this instance from outside HULAFE shall be necessary.

DFKI will provide Fitbits and a WebApp for activity and nutrition tracking. In case Smartphones are unavailable, DFKI can provide a limited amount of tablets for the participants. The WebApp will be in Spanish and be hosted on server running in DFKI premises in Germany. A permanent internet connection is necessary when using the WebApp. Data of participating volunteers will be stored under an ID on the server, together with the account names of the associated Fitbits. No personal details allowing for identification is stored on that server. DFKI will provide a secure interface to allow to setup new patients in their system, which essentially consist of assigning an ID and associated a Fitbit device.

HULAFE will have an API to pull collected nutritional and activity information of volunteers involved in the study by using the ID. Thus, HULAFE needs to maintain the relationship between the actual patients and that ID and is the only institution that knows that. The pull mechanism is secured by standard authentication and encryption mechanisms.

Using the pull mechanism, HULAFE can regularly integrate the activity and nutritional data into the holistic health records in their instance of the CrowdHEALTH platform.

The WebApp will be improved during the project and have several levels: In a first level, it will mainly help collecting nutritional data and provide statistical overviews/summaries for the persons, which can also show this to their nutritionists. Based on feedback of the participants the interface will be improved and in later stages also include suggestions of meals, first without being tailored to the health conditions. A later personalisation based on personal preferences, health data or dietary plans from nutritionists may be included, but only as suggestions without any warrantee.

For the implementation part, the selection of participants is done by the endocrinologists in the outpatient consultation. The target is to have at least 50 participants. Participation is voluntarily and based on an informed written consent. Volunteers can end their participation at any time but need to return the Fitbit devices and tablets.

When a patient/volunteer participates in the use case, then the nutritionists at HULAFE will inform them about the Fitbit and WebApp device and help them to set it up. This will be on the basis of

- An instruction document for the nutritionists
- A manual for the volunteers plus a ‘Help’ functionality inside the WebApp.

All documents will be provided in Spanish.

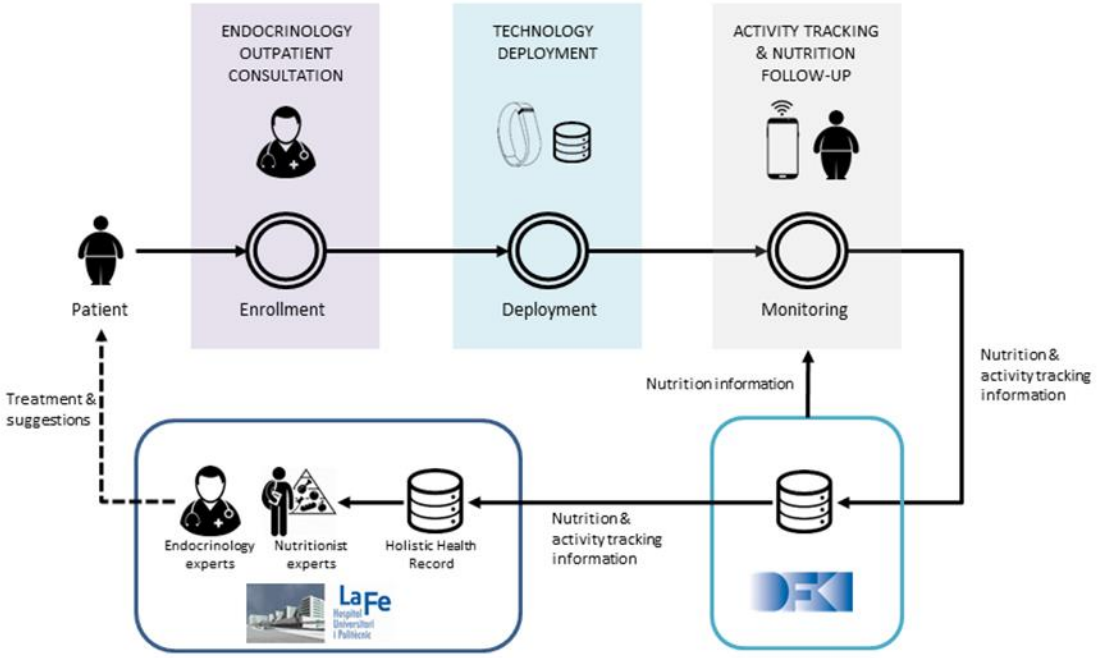

Figure 5: Schema illustrating the cycle of the pilot study of the integrated Use Cases

A prerequisite for inclusion of volunteers is that they have wireless internet available at home (and ideally also when not at home, in order to have a direct collection of nutritional data through the WebApp). In addition, they will be adult patients having overweight or obesity but without morbid obesity, that is, with a BMI between 27 and 35 kg/m2, and without any functional disability.

The nutritionists will use the DFKI API to setup the volunteer, which will create an ID in the DFKI system and link it with the Fitbit and the WebApp. That ID is then stored in HULAFE's patient health record. When a patient decides to leave the study, a similar interface is used to unlink him from the system. Part of the written informed consent shall regulate what happens to the collected data so far (if it is deleted entirely, if it can be kept in anonymized form etc).

Implementing the use case will require approval by the ethics committee in HULAFE and DFKI will provide all necessary information to prepare the necessary documents. This may also include a Data Protection Agreement between HULAFE and DFKI.

The use case shall start with collecting data in March/April 2018. Participants will be enrolled continuously based on availability. If a participant drops out, then a new participant is included as soon as possible. Ideally the participants stay with the project until the end of evaluations to collect as much data as possible. However, participating at least one year is a target lower limit in order to complete a full year with all seasons, which typically impacts on activity types and levels as well as available food and seasonal nutritional habits (festive seasons, holidays, for instance).

For the DFKI-CRA integration a certain number of existing CRA Use Case participants (cancer patients) will receive a fitness tracker each. As a result, the data already collected by CRA will be enriched with raw biodata – that is the heartrate and the step-count at any minute a day - about physical activities gathered with fitness trackers (Fitbit).

Some patients will receive a fitness tracker, which they will be able to use for at least one year. They will have full access to the information and device management abilities provided directly by the device manufacturer (e.g. mobile app, web portal).

CRA may also enable participants to enrich the gathered activity information with additional data about them. Such data can include, for instance, the exact type of physical activity (for example: jogging, swimming, football, biking…), as well as the number of people involved in the activity. This will need to be considered along two parameters: firstly, cancer patients often have limited range of physical activities they can participate in, due to the impact of the condition and its treatments. Second, engaging people in "active" tracking of activities and requiring provision of additional data has been proven quite challenging in a real-world setting; in this case, the challenges are compounded by the small number of participants, as well as the limited value they would get from describing their activities. Moreover, the ability to characterise activities will be possible by understanding the different activity patterns through the data collected by the device, and only if the patients will be able to remember what they were doing for each of the parts of the day this data represents (and for the past number of days for which they have to enter data.

This integrated Use Case will enable us to add further important parameters in the analysis of the main Use Case through the CrowdHEALTH platform.

More specifically, the following will be feasible:

- Comparison of physical tracker adherence (use) vs online coaching adherence: Some people are more likely to use a physical device rather than a purely digital service. This is because of a combination of (a) the digital service being often "out of sight thus out of mind" (b) the physical device being something they become accustomed to using without realising it. Differences in adherence will be evaluated among the patients who receive the trackers.

- Changes in online coaching adherence among patients receiving a tracker: CRA will try to provide trackers to patients who have used the online coaching service for some time. This will allow the identification of changes in behaviour towards the online coaching service.

- Comparison of adherence between patients who are given a tracker vs those who are not: Through the CrowdHEALTH platform, it will be possible to create groups of "similar" patients and compare the behaviour depending on whether they have

received a tracker or not. Such an analysis will allow the evaluation of the general impact of a tracker device on online coaching use.

- Comparison of quality of life elements between patients who are given a tracker vs those who are not. Assuming that the use of a tracker also incentivises patients to be healthier overall, the CrowdHEALTH platform will enable the analysis of quality of life impact. Furthermore, comparison of quality of life outcomes among “similar” patients will be possible, thus allowing for the construction of causality relationships between various elements.

DFKI will provide Fitbits and API endpoints to pull the gathered biodata from Fitbit. This will be the same API as HULAFE is using.

No personal details allowing for identification are stored on that server. DFKI will provide a secure interface to allow to setup new patients in their system which essentially consist of assigning an ID and associating a Fitbit device. The link between patient and ID is only known to CRA. At specific time intervals, CRA will receive the corresponding activity data from the DFKI server (asynchronously, offline), in order to enable aggregations, correlations and further processing through the CrowdHEALTH platform.

Selection and management of participants is done by CRA. The target is to have at least 30 participants. Participation is voluntary and based on an informed consent recorded digitally. Volunteers can end their participation at any time.

When a patient/volunteer participates in the use case, they will be informed by CRA about the Fitbit setup and help them to set it up. This will be on the basis of

- An instruction document for the CRA personnel
- A manual for the volunteers on how to setup their Fitbit devices.

All documents will be provided in English.

The use case shall start with collecting data in March/April 2018. Participants will be enrolled continuously based on availability. If a participant drops out, then a new participant is included as soon as possible. Ideally the participants stay with the project until the end of evaluations to collect as much data as possible. However, participating at least one year is a target lower limit in order to complete a full year with all seasons, which typically impacts on activity types (festive seasons, holidays, for instance).

3.6.2. Key Players/Actors

The key players for the use case during the project are the participating patients, the DFKI and HULAFE. DFKI will collect the data from the HULAFE patients as described above and store it in a protected, proprietary database of HHR records for privacy reasons. Aligned with national and European legal data privacy regulations, this information will be processed (e.g.,

anonymized, aggregated, etc.) to be included in the CrowdHEALTH instances of the HULAFE and other partners. Those instances of the CrowdHEALTH platform shall allow running the developed analytics and the HHRs collected in the DFKI use case and other HHR records from other partners. The policy makers identified for the project (see below for more on these) can have access to the CrowdHEALTH platform and have access to the results of the analytics.

3.6.3. Relevant Policies

Relevant policies for the DFKI Use Case are all those related to the combination of health, nutrition and physical and social activities. This may be research policies to investigate observed co-occurrences of specific habits (deficiencies, excesses, unbalance, versatility), lifestyle, environment and good or bad health. This may also be prevention policies, e.g. have population-specific more education on nutrition or physical and social activities, foster widespread availability of specific nutritional offers or support programmes for active and healthy living. The following is a list of institutions that are concerned with the aforementioned kinds of policies:

- Health insurance companies (individuals or general associations of these)
- Expert consultant to and members of ministries of health, research or education (regional, national or European)

  1. To prioritize initiatives for the population
  2. To develop research programmes

- Public television services to program thematic priorities

3.6.4. Key Performance Indicators

From the DFKI and HULAFE point of view the adherence to using the activity and nutrition tracker are the main KPIs that can be measured. Example KPIs for this are:

1. Continuity of usage of the WebApp and Fitbit
2. Acceptability and System-Usability-Scale based on respective evaluation forms from the patients' point of view as well as from the healthcare professionals' point of view.
3. Amount of feedback about errors in nutritional data, addition of new nutritional data, feature requests (more feedback indicating more involvement and thus an active usage of the system)

Any other effect on the self-awareness of the persons would be nice but difficult to measure and verify the influence of the system alone as well as on a general level due to the small size of involved participants.

From the DFKI and CRA point of view the adherence to using the activity tracker are the main KPIs that can be measured. Example KPIs for this are:

1. Physical tracker adherence (use): days/weeks/months of continued use, etc.
2. Online coaching adherence among patients using physical tracker.
3. Quality of life evaluation among patients using physical tracker.

Any other effect on the self-awareness of the persons would be nice but difficult to measure and verify the influence of the system alone as well as on a general level due to the small size of involved participants.

3.6.5. Expected Outcomes

The expected outcome is a better self-awareness for the patients and to acquire enough data and feedback to verify if the platform made a difference for the patients.

Additionally, with the aggregation of the data into the HHR platform we can provide a quick overview over the nutritional habits and the activities a user partakes in.

3.6.6. Data Flow

The sensed activity data will be collected by the Fitbit devices automatically and will be synced as often as the user syncs the device itself; those activities should be expanded by the study participant if needed. Additionally, the study participant will enter activities manually that cannot be detected from available sensor data, which will include, for instance, social activities like going to watch a movie, or taking a walk together. The data coming back from the system include aggregated sensor data for activity level, heart-rates and burned energy.

The nutritional data will be entered into the system either while cooking a recipe or if the exact recipe is unknown the system will present similar recipes based on the known ingredients. The study participant will have the possibility to estimate the nutritional data based on known recipes. The system can warn the participant if the chosen dish violates a dietary plan and can give an alternative to it and returns aggregated data about consumed energy and other nutritional information. At the end of day, the participant is asked to fill in missing information about the day, for example about his consumption of drinks of any kind or his social activities that day.

3.6.7. Contributions to HHRs

1. HHR Clinical Attributes

We will collect information about diseases or health conditions, especially those related to diets at the start of the pilot use case and will enable it to edit this data as needed, for

example if there was a screening. Additionally, also health conditions not needing direct medical attention shall be entered by the patients into the system, this may for example be a common cold.

2. HHR Sensors

We will collect the heart rate, the steps and the quality of sleep with the help of fitness trackers. This data will not have medical grade accuracy but the trackers are not disturbing the user in his daily life.

3. HHR Nutrition

The use case tries to collect all information available about nutritional intake. The user will enter what he ate based on predefined foods and recipes, and will also have the ability to add own recipes. The system then tries to find all the ingredients included and calculates the general information like energy, sugar, fat, sodium and carbs. Additionally, the composition of the food intake is also saved and can be used to warn if there are potential issues regarding diets and food intolerances.

4. HHR Stress

The stress information will be calculated for one based upon bio signals available but the use case also includes a diary like function that gathers this data on a regular basis.

5. HHR ADL

The use case will gather data about activities of daily living based upon a combination of automatic detection based on the bio signals and with user provided data about their activities.

6. HHR Life Style and Social Aspects

The socio-economic aspects are gathered on a voluntary base at the beginning of the study, but given that we will have only a small group of participants those data might not render useful.

Additionally, the activity data that was manually collected or enriched contains the number of participants attending an activity, for example how many people were playing soccer with the user, the use case will also ask the user if he ate alone or in a group because that can have an influence on the food consumed.

4. Conclusions

In this document we presented the structural description of the use case scenarios based on the information collected from the use case partners. Next, in version two we will analyse further the scenarios with the usage of UML and sequence diagrams in order to describe the relationships between the actors and the system.

Still, there is a large variety of data available from the use cases that needs to be further explored in order to examine if it is useful for the development of HHRs. Moreover, the policies creation and exploitation needs further analysis in terms of KPIs and policy making procedures. In the upcoming deliverables for the use case definition and design we will focus more on the HHRs, KPIs and final users aiming to understand further the usability and the effectiveness of CrowdHEALTH platform for the responsible stakeholders.

In particular, we will analyse the interaction and the communication between the different actors focusing on the time sequence of different actions and the alternative options. This analysis will give us the opportunity to identify the operational requirements within the classes and the functions of the system.

Finally, in the next deliverable we will analyse further the integrated use case aiming to present the holistic approach of the scenarios and the potential benefits of the synergy between the CrowdHEALTH partners. We will focus on how the synergy will be effective in terms of enhancing the HHRs and how the final outcomes will help further the policy creation, evaluation and implementation processes for public health policy makers.

5. References

[1] P. Puska, C. Nishida and D. Porter, “Obesity and Overweight,” World Health Organization, 2003.

[2] K. E. Corey and L. M. Kaplan, “Obesity and liver disease: the epidemic of the twenty-first century,” *Clin Liver Dis,* vol. 18, no. 1, pp. 1-18, 2014.

[3] T. L. Visscher, B. L. Heitmann, A. Rissanen, M. Lahti-Koski and L. Lissner, “A break in the obesity epidemic? Explained by biases or misinterpretation of the data?,” *Int J Obes (Lond),* vol. 39, no. 2, pp. 189-98, 2015.

[4] C. A. Lemay, S. Cashman, J. Savageau, K. Fletcher, R. Kinney and E. Long-Middleton, “Underdiagnosis of Obesity at a Community Health Center,” *J Am Board Fam Pract,* vol. 16, no. 1, pp. 14-21, 2003.

[5] M. R. Riley, N. M. Bass, P. Rosenthal and R. B. Merriman, “Underdiagnosis of pediatric obesity and underscreening for fatty liver disease and metabolic syndrome by pediatricians and pediatric subspecialists,” *J Pediatr,* vol. 147, no. 6, pp. 839-42, 2005.

[6] J. Ma, L. Xiao and R. S. Stafford, “Under-diagnosis of Obesity in Adults in U.S. Outpatient Settings,” *Arch Intern Med,* vol. 169, no. 3, pp. 313-4, 2009.

[7] A. I. Patel, K. A. Madsen, J. H. Maselli, M. D. Cabana, R. S. Stafford and A. L. Hersh, “Underdiagnosis of pediatric obesity during outpatient preventive care visits,” *Acad Pediatr,* vol. 10, no. 6, pp. 405-9, 2010.

[8] Socialstyrelsen, “Nationella riktlinjer för hjärtsjukvård,” Socialstyrelsen, Stockholm, 2015.