preprint of the publication:

G. Drosatos, K. Bakirlis, P.S. Efraimidis and E. Kaldoudi, Communicating Personalized Risk Factors for Lifestyle Coaching, HealthInf 2018: 11th International Conference on Health Informatics, Funchal, Madeira, Portugal, January 19-21, 2018
Communicating Personalized Risk Factors for Lifestyle Coaching

George Drosatos¹, Kyriakos Bakirlis², Pavlos S. Efraimidis² and Eleni Kaldoudi¹
¹School of Medicine, Democritus University of Thrace, Dragana, Alexandroupoli 68100, Greece
²Department of Electrical and Computer Engineering, Democritus University of Thrace, Kimmeria, Xanthi 67100, Greece
{gdrosato, kyribaki, pefraimi}@ee.duth.gr, kaldoudi@med.duth.gr

Keywords: eHealth, Personal Health Systems, Risk Factors, Healthy Behaviour, Lifestyle Related Disease, Mobile Application, Privacy by Design.

Abstract: Chronic non-communicable diseases such as diabetes, chronic cardiorenal and respiratory disease and cancer, are serious, burdensome and costly conditions that share a common characteristic: they heavily depend on common behavioural risk factors, such as physical activity, diet, stress, and substance abuse. Despite concerted efforts it has been remarkably difficult to change such lifestyle related disease determinants, as behavioural change is a complex process requiring significant personal responsibility. In this paper we propose a personal mobile eHealth application to communicate personalized lifestyle related health risks and understand their individual impact on personal health condition and disease progression.

1 INTRODUCTION

Lifestyle related diseases are multifactorial, chronic, costly diseases that create a world-wide burden on individuals, and society. However, they share a common characteristic: underlying disease determinants hugely depend on lifestyle choices and environmental factors, thus can be modified so that disease can be prevented and its progression can be halted.

Current trends in health informatics reflect an important contemporary shift towards citizen engagement for health, wellbeing and disease prevention, as opposed to disease management. However, broadcasting generic health messages (e.g. ‘do this, don’t do that’) has limited effects unless there is a convincing, easily perceived and personally customized body of evidence to back healthy choices.

In this paper we propose a novel mobile eHealth application to communicate personalized health risks and coach the user to understand their impact on various aspects of personal health and plan changes in lifestyle related risk factors. The proposed personal application consumes publicly available open data on health risk evidence via standardized programming interfaces and meshes these data with private personal health information. The output is graphical interactive representation of common health risks as personalized to each individual. Privacy is engineered into the design of the application. The ultimate goal is to provide the means to foster understanding of the complex interdependent nature of personal lifestyle choices as a disease determinant.

2 BACKGROUND

In the second global status report in a triennial series published in 2014, WHO reports that chronic non-communicable diseases such as cardiovascular disease, cancer, diabetes and chronic respiratory disease, were responsible for 38 million deaths per year (68% of all deaths) and are projected to increase from 38 million in 2012 to 52 million by 2030 (WHO, 2014).

For example, the total number of people with diabetes has risen from 108 million in 1980 to 442 million in 2014 (WHO, 2014), reaching epidemic proportions globally. Diabetes is profoundly impacted by lifestyle options including dietary intake, exercise, stress, sleep, and use of alcohol. Interventions which change each of these lifestyle behaviours are shown to improve the health of diabetic patients (Spruijt-Metz, 2014).
Another major category of lifestyle related diseases includes cardiovascular diseases, including heart attacks, strokes, coronary heart disease, cerebrovascular disease and peripheral arterial disease. Behavioural risk factors are responsible for about 80% of cardiovascular disease instances and include unhealthy diet, physical inactivity, and tobacco and alcohol use. Overall cardiovascular disease is currently estimated to cost the EU economy almost €210 billion a year, of the total cost of cardiovascular disease in the EU, 53% (€111 billion) is due to health care costs, 26% (€54 billion) to productivity losses and 21% (€45 billion) to the informal care of people with cardiovascular disease (European Heart Network, 2017).

Therefore, WHO has developed programmes to reduce lifestyle related disease incidence by limiting risk factors such as tobacco, physical inactivity, dietary factors, obesity and overweight, alcohol use, and environmental pollution. Overall, large, multicentre cohort studies (Knöops, 2004; Moreno, 2008; Haveman-Nies, 2001) have shown that even small changes in lifestyle can make an important difference towards health improvement and disease reduction.

Changing such lifestyle related disease determinants represents the single biggest opportunity to improve health outcomes while bringing costs under control. But the ‘easy stuff’ is far from easy. Despite concerted efforts by policy makers, providers and payers — not to mention the best intentions of individuals — it has been remarkably difficult to effect behavioural change (Ernst & Young, 2012).

Behavioural change is a complex process requiring significant personal responsibility. Without personalized predictive information, it is not possible to leverage on this. It is also important to understand your health risks in order to benefit from news and research about specific diseases and plan preventing monitoring.

2.1 Risk Evidence in Medicine

A health risk factor is a variable (demographic, genetic, behavioural, medical, or even environmental) which when present in an individual increases the probability of a (usually) negative outcome to occur. This risk association between the exposure agent and the outcome is conveyed via a relative risk value.

For example, medical evidence suggests that obesity is a risk factor for coronary arterial disease (Guh, 2009). In this particular systematic review and meta-analysis, men with a body mass index (BMI) between 25 and 30 kg/m² were found to have a relative risk ratio of 1.29 to develop coronary arterial disease (as compared to normal male of a BMI 18.5 to 25 kg/m²). A different risk association (for the same risk factor) is found for men of a BMI greater than 30 kg/m², who present an elevated relative risk ratio of 1.72. The same evidence source shows that women with a BMI between 25 and 30 kg/m² have a risk ratio of 1.80 to develop coronary arterial disease and when their BMI is above 30 kg/m² the risk ratio is elevated to 3.10. Thus, in this example, four different associations are described between obesity (and age) and the outcome of coronary arterial disease.

Risk factors are derived from population statistical studies. Systematic reviews select similar studies and via statistical meta-analysis combine results to improve the estimates of risk ratios and increase the level of evidence. Several risk associations are then used to build total risk estimation models for the prediction of certain outcomes that are affected by multiple risk factors at the same time. For example, in cardiovascular disease, several estimation systems exist, that vary considerably in terms of the population size and characteristics, their statistical considerations, validation, and set of risk factors considered, e.g., cardiovascular risk, include the Framingham equation (Sheridan, 2003), the Joint British Societies (JBS) formula (Boon 2014), the ASSIGN score (Woodward, 2007) and the SCORE risk charts (Perk, 2012).

Conventionally, risk factors form the basis of medical guidelines and are routinely communicated to health care professionals to formulate the basis of clinical patient management.

Evidence on risk is customarily published in medical scientific press. Indicative of current research emphasis on health risk prediction is the exponential increase of published research papers in the field. To illustrate this, we have placed a crude PubMed query on risk prediction as follows:

To derive the total number of PubMed publications, we placed the following query:

"0001/01/01"[PDAT]:
"2016/12/31"[PDAT]

Figure 1 shows the number of retrieved publications per year plotted as a percentage of the total number of PubMed indexed publications and as an absolute number. The query retrieved more than 460,000 related publications that amount to almost 2% of the total PubMed indexed publications. Overall, there is an increase in the absolute number of retrieved publications over the years. When seen as the percentage of the overall PubMed corpus growth, linear regression for the last 40 years shows a statistically significant increasing trend (regression coefficient = 0.000898, p-value < 0.001, R-squared = 98.5%).

A number of models have been proposed for capturing various aspects of the clinical research process that aims to generate new risk evidence. Examples include the Ontology-Based eXtensible data model (OBX) (Kong, 2011) and the Ontology of Clinical Research (OCRe) (Sim, 2014); both aim to support the process of generating new scientific knowledge in medicine, rather than actual evidence on risk prediction. Scientific knowledge on risk factors is captured by CARRE risk factor ontology (Third, 2015). This ontology has been used to create an open data repository of structured information on risk factor evidence in the medical domain of cardiorenal disease and comorbidities, as part of a European Commission funded project CARRE (FP7-ICT-611140). The repository was developed by medical experts following a methodology based on selecting recent highest evidence sources (meta-analyses where available) and allowing for peer review of manually entered risk factor data. This RDF repository is freely available via a SPARQL endpoint, and currently holds structured descriptions of more than 100 risk factors, corresponding to a total of 250 different risk associations between more than 50 medical conditions.

2.2 Risk Prediction and Patient Empowerment

Recent years have seen a paradigm shift towards patient-centred care and health management. Under this view, patient empowerment has emerged as an approach that can help improve medical outcomes while lowering costs of treatment by facilitating self-directed behaviour change. The concept seems particularly promising in the prevention and management of chronic diseases, and it is directly connected with personalized patient services, education and preventive measures.

At the First European Conference on Patient Empowerment held in Copenhagen, Denmark on 11–12 April 2012 (WHO, 2012), R. Johnstone of the International Alliance of Patient Organizations claimed that: "What needs to happen is for doctors to come down off their pedestal and for patients to get up off their knees". Currently, the research community appreciates that improving a person's ability to understand and manage his or her own health and disease, negotiate with different teams of health professionals, and navigate the complexities of health systems is crucial to achieving better health outcomes (The Lancet Editor, 2012).

A recent pilot study deployed a self-monitoring intervention for cardiorenal patients coupled with rich health risk visualizations as derived from the CARRE repository (Zhao, 2017). The intervention requires the patient to use a set of personal and wearable sensors to monitor health status and couples these real-time data with medical information retrieved from personal health records. Personal private data are then coupled with risk evidence to create medical alerts for the management and prevention of disease progression in cardiorenal patients or those at risk of this disease.
A randomized controlled trial (Kizlaitis, 2016) showed that the intervention statistically increased empowerment by a 12.4% in metabolic syndrome patients, and health literacy by 21.3% in patients with heart failure of chronic kidney disease.

3 COMMUNICATING PERSONALIZED HEALTH RISKS

In this paper we propose a personal eHealth smartphone application to communicate health risks to healthy citizens or people at increased risk of chronic disease. The aim is to provide a personalized assessment of lifestyle related risks that create the necessary knowledge background for people to make educated choices towards healthier lifestyle adjustments. To achieve this goal, the Risk Coach application merges open data on health risk evidence with private health related information.

The core of the application is the current evidence on health risks. Such evidence is extracted from the public CARRE risk factor repository available at [https://devices.duth.carre-project.eu/sparql](https://devices.duth.carre-project.eu/sparql). Data on risk evidence is in a structured RDF format and contains information on exposure and outcome, risk ratio value (with confidence intervals) and the conditions under which this risk ratio value is valid. This condition is based on the characteristics of the population group used to study the particular risk factor and also includes levels of exposure. For example, the repository contains evidence on obesity as a risk factor for cholelithiasis as reported in a recent meta-analysis (Guh, 2009). Evidence on this risk factor reveals four different risk associations, each corresponding to a different condition and presenting different risk ratio values (Table 1). In this example, the condition under which obesity is a risk factor for cholelithiasis is a combination of the sex and the value of body mass index (BMI). It is evident that obese females (body mass index ≥ 30) are at an increased risk of cholelithiasis (relative risk of 2.32) as compared to overweight females (25 ≤ body mass index < 30) and as compared to obese males, who appear to have about the same risk as overweight females. The exact form of the condition (the variables and their values) depends on the clinical study and its population groups. In fact, risk associations in CARRE repository at present use a total of 95 different variables to formulate conditions, and this number will most likely increase as new risk factors are entered in the repository.

<table>
<thead>
<tr>
<th>Description fields</th>
<th>Risk association #1</th>
<th>Risk association #2</th>
<th>Risk association #3</th>
<th>Risk association #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk factor:</td>
<td>obesity</td>
<td>obesity</td>
<td>obesity</td>
<td>obesity</td>
</tr>
<tr>
<td></td>
<td>[is an issue in]</td>
<td>[is an issue in]</td>
<td>[is an issue in]</td>
<td>[is an issue in]</td>
</tr>
<tr>
<td></td>
<td>cholelithiasis</td>
<td>cholelithiasis</td>
<td>cholelithiasis</td>
<td>cholelithiasis</td>
</tr>
<tr>
<td>Condition:</td>
<td>body mass index ≥ 25 AND body mass index &lt; 30</td>
<td>body mass index ≥ 30 AND sex = 'male'</td>
<td>(body mass index ≥ 25 AND body mass index &lt; 30) AND sex = 'female'</td>
<td>body mass index ≥ 30 AND sex = 'female'</td>
</tr>
<tr>
<td></td>
<td>AND sex = 'male'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratio type:</td>
<td>relative risk</td>
<td>relative risk</td>
<td>relative risk</td>
<td>relative risk</td>
</tr>
<tr>
<td>Ratio value:</td>
<td>1.09</td>
<td>1.43</td>
<td>1.44</td>
<td>2.32</td>
</tr>
<tr>
<td>Confidence Interval</td>
<td>0.87 – 1.37</td>
<td>1.04 – 1.96</td>
<td>1.05 – 1.98</td>
<td>1.17 – 4.57</td>
</tr>
<tr>
<td>PubMed ID:</td>
<td>19320986</td>
<td>19320986</td>
<td>19320986</td>
<td>19320986</td>
</tr>
</tbody>
</table>

In order to personalize risk communication, the appropriate personal information needs to be known and compared to the known risk factor conditions in the public repository. This is ensured via a Risk Coach application module that allows the user to input appropriate personal information. To increase efficacy and reduce privacy concerns, the application dynamically recognizes different variables available in the CARRE risk factor repository and actively builds the personal information input fields at real time.

By default, the user is required to input some basic demographic and biometric information such as age, sex, height and weight. This initial information is used to calculate any personal risks, if these apply based on the specific values describing the user. To investigate whether further risk factors apply for the particular user, the application queries the external risk factor repository for all known
variables and dynamically constructs the appropriate fields for the user to fill in personal values. Personal risks are calculated by checking which of the risk association conditions available in the external database are valid for the user inserted personal data.

### 3.1 An Architecture to Preserve Privacy by Design

The *Risk Coach* application is designed to be used by the citizens for self-management of health and disease mainly outside the formal healthcare context. Based on the requirements for personal data communication and an analysis of privacy issues in personal e-health systems (Drosatos, 2016), we can identify the following functionalities that may introduce privacy issues in Risk Coach application:

1. personal data storage and processing;
2. integration of personalized public data.

When both storage and processing of personal data are located on a user personal device, privacy can generally be achieved by default. This approach is followed in Risk Coach where personal data and personal risk calculations are stored locally on the application. Also, all risk calculation is performed locally on the smartphone device, as shown in Figure 2 which presents an overview of the application architecture.

![Figure 2: An overview of the Risk Coach mobile application functional modules and architecture.](image)

A potentially important and usually elusive privacy issues may arise when at run time public data are requested from external sources (in this case the risk factor repository). Although the data are publicly available, just the act of linking particular data to a specific user may cause a privacy violation, by revealing the user’s health condition. There are a number of proposed techniques to conceal user requirements by altering the initial request, e.g. by expanding and generalizing the request for public data. These techniques fetch a large amount of data to the user application and then a second round of local processing extracts the specific data relevant to the user (Efraimidis, 2016). Other emerging approaches require the cooperation of a group of users in the system to conceal one another’s requests (Romero-Tris, 2015). An alternative is to use anonymous network technologies that protect the physical address of user from the public service. A representative example is the TOR service (Dingledine, 2004), which creates a network of proxies over the internet and allows recursive message encryption along the chain of proxies. In the current approach, the entire risk factor dataset is locally retrieved, so as to hide user specific requests.

### 3.2 Implementation

The *Risk Coach* application is currently implemented for Android mobile devices using Android Studio (version 2.3.3) and it is compiled for Android 7.1 (API level 25) with backward compatibility until the Android 4.4w (API Level 20). Animated graphs are produced using the MPAndroidChart library ([https://github.com/PhilJay/MPAndroidChart](https://github.com/PhilJay/MPAndroidChart)) and the retrieval of data
throw the CARRE SPAQL Endpoint is performed using the OkHttp library (https://square.github.io/okhttp/).

The application supports the following main functionalities (implemented as individual Android application fragments): (a) visualize the generic risk factor evidence available in the public CARRE risk factor repository, as an overall repository summary and browser; (b) create of a local copy personal health related information, with fields dynamically compiled based on what information is required to customize available evidence on risk associations; and (c) view and interact with personalized risk information.

Indicative screenshots are shown in Figure 3. Personalized risk visualization allows the user to initially explore a summary of personal risk factors. Each risk can then be expanded in a graph chart to show all potential outcomes and the particular risk ratio information. The user can also interact with personal health information to set lifestyle goals and observe the expected results in terms of risk reduction; for example increase or decrease weight to realize its impact on obesity related risk factors.

![Figure 3: Screenshots of Risk Coach mobile Android application.](image)

4 DISCUSSION AND FURTHER WORK

Patient empowerment is about designing and delivering health and social care services in a way that they are inclusive and enable citizens to take control of their health care needs. According to the European Network for Patient Empowerment (ENOPE), the first mandate for an activated patient is to be able to understand their health condition and its effect on their body.

The Risk Coach application presented in this paper addresses directly this basic requirement by communicating state-of-the-art medical evidence on common lifestyle related health risk determinants as personalized to the individual. The application enables the individual to have an overall view of his/her own health status; understand potential health deterioration and disease progression based on current health status; and visually and quantitatively investigate the impact each health determinant may have on various health conditions.

Work in progress involves a randomized controlled trial to access the efficacy of communicating health risks to the general public via the personalized application that preserves privacy by design. The aim of the study is to assess user satisfaction and efficacy of the application to empower people and coach them towards a healthier lifestyle. Primary objectives are to increase health literacy, and increase level of empowerment. Secondary objectives are to improve lifestyle habits (smoking, physical activity, adherence to self-
monitoring and therapy) and test for application acceptability and user satisfaction. The study considers two different population groups: healthy volunteers and chronic heart disease patients. The control arm will be given general written information on common lifestyle related risks; the intervention arm will be given access to the Risk Coach mobile application.

Health literacy, before and after intervention, will be assessed via the European Health Literacy Questionnaire (Sorensen, 2013) and the Lipkus Expanded Health Numeracy Scale (Lipkus, 2001). Patient empowerment will be assessed via the SUSTAINS instrument (Unver, 2013) which considers three axes: enabling and strengthening empowerment of patients; enabling better medical results; and enabling a more efficient use of healthcare resources and containing costs. Finally, following the ISO standard ISO 9241 Part 11, system usability will be measured by taking into account the context of use of the system — i.e., who is using the system, what they are using it for, and the environment in which they are using it. Along with an informative system evaluation, we will also deploy the System Usability Score instrument (Brooke, 1996).

Communicating health risks via interactive personalized risk visualizations is expected to increase health awareness, motivate persons to adopt a healthier lifestyle, and contribute towards increasing public health literacy and informed shared decision making.

ACKNOWLEDGEMENTS

The work presented in this paper was partly sponsored by the FP7-ICT project CARRE (Grant No. 611140), funded in part by the European Commission and Greek National Matching funds.

REFERENCES

Ernst & Young, 2012. Global Life Sciences Report, Global Lifesciences Center Ltd.


