Using non-medical risk factors related to dementia and cognitive decline for developing an evidence-based e-health tool

ANUSHARANI GOPU
Using non-medical risk factors related to dementia and cognitive decline for developing an evidence-based e-health tool

By

Anusharani Gopu

Master thesis submitted in partial fulfilment of the requirement for the degree of Master of Science in Distributed Systems.
(School of ICT, KTH)

Supervisor: Marie Sjölinder
Examiner: Magnus Boman
Acknowledgement

First and foremost, I would like to thank my examiner Prof Magnus Boman, without whom it wouldn’t be possible to complete this thesis work. He was kind and patiently answered all my queries without any delays and was accessible throughout the summer when I did most of the work. I also would like to thank my supervisor Marie Sjölinde for her input and valuable suggestions mainly regarding the structure, integrity and quality of the report. I am also thankful to Caroline Wärn, Mattias Jacobsson and all other people at SICS for their valuable input and support.

Thanks to Shireen Sindi and Rui Wang at Aging Research Centre (ARC), Karolinska Institute, for their valuable time in explaining risk score calculations and statistical modelling. Special thanks to my employer Bo Öström at Bogleboo AB who granted me leave from my duties to perform this thesis work. Last but not least, I would like to thank my family for giving me the moral support during the thesis work.

Anusharani Gopu

KTH, September 2016
Abstract

The number of dementia cases is increasing worldwide. Most research and development in this area is related to the prevention of dementia, and to the development of various prediction tools for dementia. The tools made available take most of the medical data into account while calculating risk scores, with only a small amount of non-medical data. There is a lot of data related to medical and non-medical risk factors available from various sources which can be retrieved and analysed in real time, but this is today not used in any risk score tool for risk score calculation. As part of the project Multimodal strategies to promote a healthy brain in ageing: Innovative evidence-based tools (MULTI-MODE), a new risk score is being developed to be used in a new ICT-based tool for dementia prediction. Identification of non-medical data and a good model to fill the gap between data available at the server and using this data in risk score calculation may help in increasing the predictability of tools. In this thesis, some of the existing risk factors for the prediction of dementia are described, and the importance of non-medical factors in calculating risk scores is discussed. Additional non-medical factors are identified that could be included in future versions of the risk score. A database design for storing risk score information efficiently is presented, as is an app structure that can be used at the server side to validate the user input and to increase the effectiveness of a prediction tool.

Keywords

Alzheimer’s disease, Dementia, Mild cognitive impairment, Risk factors, Dementia risk score, Cognitive decline, Cognitive training, Dementia prevention, Evidence-based prediction, Risk score app.
Sammanfattning


Nyckelord

Alzheimers sjukdom, Demens, Kognitiv svikt, Riskfaktorer, Kognitiv träning, Demensförebyggande, Evidensbaserad förutsägelse.
# Table of contents

1  Introduction .................................................................................................................. 1
   1.1  Problem .................................................................................................................... 2
   1.2  Purpose ..................................................................................................................... 2
   1.3  Goal .......................................................................................................................... 3
   1.4  Methodology ............................................................................................................ 3
   1.5  Thesis overview ........................................................................................................ 4

2  The role of ICT in active ageing and risk factor categorisation .................................... 5
   2.1  The role of ICT in active ageing ............................................................................. 5
   2.2  The role of mobile apps in the health sector ............................................................ 6
   2.3  Role of risk factors in risk calculation ..................................................................... 6

3  Targeting the individual ................................................................................................... 8
   3.1  Personal genomics ..................................................................................................... 8
   3.2  Personalised health .................................................................................................... 10

4  Risk scores ....................................................................................................................... 13
   4.1  FINGER ................................................................................................................... 13
   4.2  CAIDE risk score ..................................................................................................... 14
   4.3  Australian National University AD Risk Index (ANU-ADRI) .................................... 17
   4.4  Comparison between old CAIDE, ANU-ADRI and new CAIDE risk scores: .......... 18
   4.5  Late Onset Alzheimer’s disease risk score ............................................................... 19
   4.6  Late-life dementia risk index .................................................................................. 20
   4.7  MULTI-MODE .......................................................................................................... 21

5  Non-medical risk factors .................................................................................................... 24
   5.1  Geographical related risk factors ............................................................................ 25
       5.1.1  Air pollution ....................................................................................................... 25
       5.1.2  Geography (Area of living/Ethnicity) ................................................................. 25
   5.2  Psychological risk factors ....................................................................................... 25
       5.2.1  Developmental and early-life risk factors / Stress level .................................... 25
       5.2.2  Depression ........................................................................................................ 26
       5.2.3  Sleep deprivation .............................................................................................. 26
   5.3  Lifestyle-related risk factors ..................................................................................... 26
       5.3.1  Smoking ............................................................................................................. 26
       5.3.2  Alcohol consumption ....................................................................................... 27
       5.3.3  Mid-life coffee and tea drinking and the risk of late-life dementia .................... 27
       5.3.4  Social isolation at old age .................................................................................. 27
       5.3.5  Lack of engagement in reading and hobbies ...................................................... 27
6 Analysis ................................................................................................................................. 28
  6.1 Old CAIDE risk score algorithm ....................................................................................... 28
  6.2 Introducing non-medical risk factors into risk score ......................................................... 29
  6.3 Server-side data ............................................................................................................... 32
  6.4 Database design ............................................................................................................... 33
      6.4.1 To store client (patient) information ........................................................................ 33
      6.4.2 Caregiver information .............................................................................................. 34
      6.4.3 Device information .................................................................................................... 35
      6.4.4 Risk factors ............................................................................................................... 36
      6.4.5 Risk score information .............................................................................................. 36
  7 Conclusion ............................................................................................................................. 37
      7.1 Discussion ...................................................................................................................... 37
      7.2 Concluding remarks ..................................................................................................... 40
List of figures

Figure 1 Venn diagram representing AD as subset of Dementia, Dementia as subset of Mild cognitive impairment (MCI)........................................................................................................4
Figure 2 Classification of risk factors related to dementia. ................................................................. 7
Figure 3 Cost per Genome. .................................................................................................................... 10
Figure 4 Percentage of patient population where a particular drug in a class is ineffective........ 11
Figure 5 CAIDE Risk calculator Personal Data. .................................................................................. 16
Figure 6 CAIDE Risk score calculator level indicator graphs. ........................................................... 16
Figure 7 CAIDE Risk score calculator information pages................................................................. 17
Figure 8 Algorithm for the old CAIDE risk score............................................................................. 28
Figure 9 Suggested Object/Class diagram for risk score app development including the relationships between different objects......................................................................................... 33
Figure 10 Database relationship diagram for storing user's basic information at server............... 34
Figure 11 Database relationship diagram for storing caregiver information at server. ............... 35
Figure 12 Database relationship diagram for storing User's device information at server.......... 35
Figure 13 Database relationship diagram for storing user's risk factors with different timestamps at server..................................................................................................................................................... 36
Figure 14 Database relationship diagram showing relations between user, caregiver and users calculated risk score with different timestamps......................................................................................... 36
Figure 15 Sequence diagram: Retrieval of DBP from Caregivers Database for Risk score calculation. ......................................................................................................................................................... 37
List of tables

Table 1 CAIDE Risk Score ........................................................................................................................................................................ 14
Table 2 Comparison of risk scores .................................................................................................................................................................. 18
Table 3 LOAD Risk score .............................................................................................................................................................................. 20
Table 4 Late-life dementia risk score .......................................................................................................................................................... 21
Table 5 Example data entered by a fictive person into Risk App ............................................................................................................. 29
Table 6 Risk scores for fictive data in table 5 ............................................................................................................................................. 29
Table 7 Example data for Air Quality Index ............................................................................................................................................ 30
Table 8 Example Javascript function which demonstrates the calculation of user’s age using Swedish personal number or date of birth of the user .......................................................................................................................... 32
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and communications technology</td>
</tr>
<tr>
<td>KI/ARC</td>
<td>Karolinska Institutet/Aging Research Centre</td>
</tr>
<tr>
<td>CAIDE</td>
<td>Cardiovascular Risk Factors, Ageing and Dementia</td>
</tr>
<tr>
<td>SNAC-K</td>
<td>Swedish National study on Aging and Care in Kungsholmen</td>
</tr>
<tr>
<td>SICS</td>
<td>Swedish Institute of Computer Science</td>
</tr>
<tr>
<td>RISE</td>
<td>Research Institutes of Sweden</td>
</tr>
<tr>
<td>MULTI-MODE</td>
<td>Multimodal strategies to promote a healthy brain in ageing: Innovative evidence-based tools</td>
</tr>
<tr>
<td>EIT</td>
<td>European Institute of Innovation and Technology</td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>FINGER</td>
<td>Finnish Geriatric Intervention Study</td>
</tr>
<tr>
<td>NTB</td>
<td>Neuropsychological Test Battery</td>
</tr>
<tr>
<td>ANU-ADRI</td>
<td>Australian National University AD Risk Index</td>
</tr>
<tr>
<td>SBP</td>
<td>Systolic Blood Pressure</td>
</tr>
<tr>
<td>DBP</td>
<td>Diastolic Blood Pressure</td>
</tr>
<tr>
<td>IoT</td>
<td>Internet of Things</td>
</tr>
<tr>
<td>RCT</td>
<td>Randomised control trial</td>
</tr>
<tr>
<td>AQI</td>
<td>Air Quality Index</td>
</tr>
<tr>
<td>My-AHA</td>
<td>My Active Healthy Ageing</td>
</tr>
<tr>
<td>SNP</td>
<td>Single-Nucleotide Polymorphism</td>
</tr>
</tbody>
</table>
1 Introduction

This study is of risk factors related to dementia, particularly on non-medical risk factors that can be considered while developing evidence-based tools for predicting dementia risk. A risk factor is something that increases the likelihood of developing a condition or disease.

Dementia is caused by neurodegeneration and is characterised by a disturbance of multiple brain functions, including thinking, memory, calculation, orientation; language, learning capacity, and judgement which eventually make it difficult for people to perform daily activities [1]. Alzheimer’s is the most common type of dementia and possibly contributes to 60-70% of cases [2]. Other types of dementia are vascular dementia (e.g., strokes, heart failure), dementia with Lewy bodies (e.g., sleep behaviour disorder), and frontotemporal dementia (e.g., language disturbances, deterioration in behaviour and personality) [3].

Brain-related health issues at late-life will burden societies and family members physically, psychologically, and economically. There are around 34 million individuals worldwide suffering from Alzheimer’s disease (AD) and there is a prediction that this number will be tripled in the coming 40 years, due to changes related to demography and an increase in life expectancy [4]. Along with memory loss, Alzheimer’s patients can experience other symptoms like difficulty in judgement, thinking, and confusion related to place and time. In some cases, patients cannot carry out basic functions like swallowing and walking. This increase in dementia at older age requires more attention towards elderly people from formal caregivers (community care professionals) or from informal caregivers (friends or relatives), which also increases society health care costs tremendously [5] that contributes to one-third of all health related expenses. Costs related to dementia and AD in Europe were estimated to > €170 billion in 2008 [6]. Governments and caregiving institutions have become aware of the impact of brain problems in at-risk elderly people towards family members, societies, and individuals [7], prompting more research into decreasing future health costs through prevention. Such prevention is possible with diagnosing and treating these diseases at early stages. Information and communications technology (ICT) can play a very important role in the prediction of these ageing problems by developing applications (related to measuring diet, exercise, weight loss and fitness), tools, and sensors which can be used for early detection of these diseases [8].

Dementia impact has received the greater attention of governments and politicians all over the world in recent years. Governments of developed countries like UK, Norway, Sweden, France, U.S., and South Korea have developed specific plans or strategies [9]. Dementia cases are usually concerned with elderly people but there is an increase in early-onset dementia (under 65 years of age). Although there is an increase in early-onset dementia cases, dementia is a condition which affects elderly people and is the leading contributor to disability and dependence [10], [11]. Rapid increases in the numbers of older people are forecast for China, India, and Latin America. Dementia awareness and a health system that can tackle the dementia related problems are much more limited in less developed regions [12]. It is therefore very important to track the global pervasiveness of this difficult situation and to provide findings related to dementia. As there are no curative treatments for dementia, prevention has been highlighted as the major health priority according to the G8 dementia summit and World Health Organization [13], [14]. Intervention towards these modifying risk factors helps in delaying the onset of AD [15], and it has been estimated that 10-25% reduction in key risk factors could prevent 1.1-3.0 million AD cases internationally [15]. Therefore, it is helpful to have evidence-
based tools that measure individual’s risk factor more accurately and thereby start the treatment earliest in order to prevent the dementia progress. Prediction of risk analysis at early stages is helpful for both patients and their caregivers as they get to benefit from counselling on how to handle the diseases, and so can they plan. Early diagnosis is also helpful for proper medication.

To investigate the effect of risk factors on problems that occur in late life like dementia and cognition, the Department of Neurology, University of Finland, the National Institute of Health and welfare, Helsinki, and KI/ARC (Karolinska Institutet/Aging Research Centre) together started a research called CAIDE (Cardiovascular Risk Factors, Ageing and Dementia) and identified a number of factors that may increase or decrease the risk level of dementia or cognitive impairment. These risk factors were used to develop CAIDE risk score. The CAIDE Dementia Risk Score tool has been developed to predict the chance of occurrence of dementia in a person in next 20 years [16]. This tool was developed based on considering various risk factors like age, hypertension, physical activity, obesity, and educational background. The CAIDE dementia Risk Score tool has been validated in a multi-ethnic population in the U.S. In a study made by ANU-ADRI, the CAIDE risk score was calculated and compared in three different older age cohorts, and validation of result was found to be poor [17]. This was likely due to CAIDE development having been done by considering mid-life risk factors and validated in older cohorts. This, in turn, suggests that risk factor effects can be different in different cohorts. The CAIDE risk score was developed by considering midlife factors which may not be suitable for older cohorts. There is thus a need for a new risk score for predicting dementia in older cohorts.

Researchers at KI/ARC have developed an algorithm for a new risk score that can be applicable for the older population, based on data collected from Swedish National study on Aging and Care in Kungsholmen [SNAC-K] project [18]. As a part of this project, researchers at Swedish Institute of Computer Science (SICS) are going to produce evidence-based ICT tools to predict dementia risk and prevent cognitive decline/dementia, based on the new risk score. These tools are intended to be used by citizens as well as by healthcare staff. The main purpose of developing these tools is to reach more users, reducing health costs, social burden, and to provide better tools for dementia prediction.

1.1 Problem

Developed risk scores for dementia take mostly medical data (ex. blood pressure, cholesterol), and only a limited amount of non-medical data (ex. age, gender, diet, smoking), into account when calculating individual risk, whereas excluded non-medical data may in some cases play a significant role, and could arguably contribute to the risk score.

1.2 Purpose

There is an opportunity for ICT-based tools to take non-medical data into account that may help in developing better predictive ICT tools for dementia. The purpose is to identify such non-medical data that can be included in future risk score calculation, and a suggested database design for storing risk score information. The study can be considered a complement to the MULTI-MODE (Multimodal strategies to promote a healthy brain in ageing: Innovative evidence-based tools) project agenda, in which researchers will develop and validate innovative ICT tools, notably an app, to help predict risk scores for dementia.
1.3 Goal

The goal is to contribute to an increase in the usefulness of ICT tools that predict cognitive decline, in part through the identification of non-medical data that can be included in risk score calculation, enhancing risk score by checking the user’s input accuracy, and in part through a suggested database design for storing risk score information efficiently.

1.4 Methodology

The SICS, as one of the partners in MULTI-MODE, and a part of Research Institute of Sweden (RISE), was assigned to develop and test eHealth tools to promote healthy brain ageing and widely implement the risk scores. The thesis work was carried out at SICS. Interaction with SICS researchers part of MULTI-MODE was instrumental to this work, including participation in internal meetings and an internal workshop in May 2016. Some advice and communication with KI/ARC researchers were also helpful to obtaining the results.

As part of MULTI-MODE, tools will be designed for use by citizens and health care staff, making use of state-of-the-art mobile technology. The tools will be based on the new risk score derived from CAIDE Dementia Risk and will be validated in different datasets. The datasets are from many longitudinal cohort studies like SNAC, KI, and CAIDE, Rotterdam study, United Kingdom Clinical Practice Research Datalink and CHARIOT register. 8000 Swedish citizens participated in SNAC study, 2000 participated in CAIDE, and CHARIOT register involved 26,000 individuals. Apart from this data, risk score also considered the data from medical records, quality registers (ex. workshops, reviews, inspections, and audits), and prescribed drug registers. Developed new risk score has also used the data from H2020 Athlos project which has data related to more than three hundred thousand individuals.

At the beginning stage of the thesis work, a basic understanding of the problem background and the former research was achieved through the documents provided by Professor Magnus Boman at SICS who is the SICS project leader for MULTI-MODE. As the field of dementia risk assessment is relatively mature, a deductive literature study and desktop research were completed by collecting information regarding risk scores and other factors related to problems that occur in late life. Interviews have been conducted with KI/ARC researchers to establish the process of developing the first CAIDE risk score, which was based on mid-life risk factors. These interviews helped in understanding in depth the existing risk score but still not able to find the algorithm used in OLD CAIDE risk score. So an attempt made to produce an algorithm used for old CAIDE risk score in Chapter 6.1. For knowing about a newly developed risk score, a semi-structured interview has been conducted with Rui Wang, PhD at ARC. After studying and understanding existing risk scores it has been observed that the majority of the risk factors involved in the risk score calculation are medical-related risk factors. There is relatively little importance given to non-medical risk factors. Therefore, further study of non-medical factors was motivated. A reductionist approach has been used for identifying some additional risk factors, in which risk factors are categorised into two major categories which are further divided into sub-categories. The scope of the thesis is limited to the discussion on some of the non-medical factors instead of all identified factors and some assumptions were made about the risk factors data during the analysis. A qualitative approach has been used for understanding and identifying underlying non-medical risk factors.
In this thesis, AD refers to studies associated only with Alzheimer’s, if a study relates to serious cognitive decline in general then it is considered as dementia, and other studies related to cognitive impairments beyond those expected based on the age and education levels, but which does not affect daily life are considered as mild cognitive impairment (MCI) [19]. MCI can be seen as an early symptom of AD but it is not dementia as it does not considerably affect daily life [20].

1.5 Thesis overview

Chapter 2 describes the role of ICT in helping elderly people, an increase in usage of mobile apps in the health sector, and various risk factor categories for dementia. Chapter 3 describes the new technologies developed within ICT and medical informatics, in the context of genomics, i.e. genetic risk factors. The chapter includes the benefits of personal genomics, precision medicine, and also provides a discussion about current trends in personal genomics. This chapter also describes the possibilities of introducing genome data in risk score applications. Chapter 4 includes a discussion of different risk scores and analyses of these risk scores. Here, the CAIDE risk score is experimented with on some specific user data to understand how the existing score appears for users. The chapter describes the MULTI-MODE project that is currently undertaken within EIT (European Institute of Innovation and Technology) Health. Chapter 5 describes the additional non-medical factors which may help in increasing the effectiveness of risk score. The effect of considering some of the identified non-medical factors into the risk score, a suggested database design for storing risk score information, and an app architecture is discussed in chapter 6. Chapter 7 contains a discussion on app-related issues, followed by conclusions.
2 The role of ICT in active ageing and risk factor categorisation

Dementia is a condition that affects mostly elderly people. It may affect younger people and dementia in younger people is called as early-onset dementia, but early-onset dementia cases are few when compared to late-onset dementia. Ageing of the population is growing very fast in both developed and developing countries. Europe is one of the major regions of the world where ageing of the population is advanced [21]. Population ageing occurs when the median age rises and shifts the distribution of a country’s population towards elderly people. The median age of the population is high in Europe when compared to other parts of the world. Approximately 12.3 percent of the global population is people aged over 60 years and this may increase to 22 percent by 2050 [21]. This is going to have a dramatic effect on all aspects of society, especially healthcare. This increase in the ageing population may be a result of a decrease in fertility rate and increase in life expectancy. Monetary costs related to dementia represent a great financial burden to society. Furthermore, care towards the dementia patients is frequently associated with physical, social, and psychological stress [22]. So prevention is the best way to reduce these care costs to a greater extent, and as there are no curative treatments for dementia, prevention has been highlighted as the major health priority [13], [14]. ICT helps in prevention and better treatments for diseases. Following section explains how collaborating ICT with health helps in better elderly one’s life.

2.1 The role of ICT in active ageing

“ICT can help older individuals to improve the quality of life” according to “EU action plan on ICT and Ageing” [23 page 50]. Examples of ICT include using an app for improving diet which logs and monitors intake so that users can get regular advice about their dietary condition. One can make use of ICT to support cognitive deficiencies using sensors and actuators which can be helpful for elderly people by providing alternatives to their physical actions, such as moving or changing the position of objects (from closing doors to turning lights on and off), motion sensors, which helps in tracking a demented person. The concept of Internet of Things (IoT) (connected smart devices which communicate with each other and with software running on the cloud) also helps in taking actions, such as controlling heating and air conditioning, locking doors and windows, and giving reminders about medication. IoT technology securely collects and analyses the data from sensors and other devices. This information can be accessed by family members, healthcare professionals, or emergency services. Emergency services use the collected information, to help in emergency situations like falling of room temperature below a safe threshold which can be a life-threatening situation. This technology is very helpful for family members who live far away from elderly relatives. The policy makers of current health care systems are not only looking for efficient care solutions, but also for changing its focus from care to prevention to reduce care costs, burden on caregivers/family members, and burden on society. As part of this, current health care system focusing on preventing dementia with help of ICT by developing e-health tools for prediction of the brain related risk factors in early stages so that making these tools widely used by citizens and healthcare staff. EIT-Health and My-AHA (My Active Healthy Ageing) are the initiatives of EU which are both part of the Horizon 2020 effort, to provide good health and healthy active ageing in adults. In My-AHA, existing platforms (i.e., sensors to detect physical actives and apps to store this data) which already are owned by My-AHA partners will be integrated for the collection of data in large-scale and analyse the collected data using existing software or by improved firmware and
software. The type of data includes data related to sleep patterns, physical activities, diet, social activities, emotions, and clinical data such as blood pressure, sugar levels, Body Mass Index (BMI) etc. The data analysis results will be used by My-AHA for improving the activities and thereby reduce the “frailty risk” in older adults [24]. Apart from this Gerontechnology is one of the ICT technology which is devoted to the design of technology and environments for a better quality of life of older adults [8]. More information on Gerontechnology can be found in Appendix A.

2.2 The role of mobile apps in the health sector

In last decade the mobile app usage has increased tremendously. Mobile devices are easy to carry and are available at user’s fingertips whenever and wherever they need them. According to a recent report, in 2015 there were 120 000 mHealth apps (mHealth is the term used for the practice of medicine and public health supported by medical devices) that were available for download on AppStore. Out of 100 most downloaded apps from all categories, 33% apps were mHealth related. The report also mentions that most of these apps are very successful in terms of user satisfaction. mHealth apps are very popular right now especially in Sweden and they contribute to around 5.1% of total app downloads [25].

In order to succeed with an app, the app should meet the primary aim for which it is designed and developed for; all other extra features are nice to have in but should not affect the primary purpose. An example is “Whatsapp” which is a simple texting app where users can be in touch with their family and friends. Despite many other advanced messengers, video chats and social media sites, this app is a huge success as it has very simple and easy GUI which meets the primary purpose i.e. sending out text messages very reliably and efficiently.

In the same manner, if designed properly, mHealth apps can be very useful for the elderly population by introducing them to physical and social activities and guiding them to overcome dementia-related problems. The apps can be helpful even for the caregivers to follow their patients in an efficient way and thereby increase their lifespan. A new mobile app is being developed as part of MULTI-MODE project that shows the dementia risk prediction for the elderly people.

2.3 Role of risk factors in risk calculation

Some people have a higher risk of developing dementia while others have a lower risk of developing dementia. Dementia risk assessment tools are helpful for the users for estimating the risk of being demented in future. The risk assessment tools use risk score calculation that is developed based on risk factor data from various studies. Understanding of risk factors related to disease may helpful in the prevention of disease. A risk factor is anything that increases the risk of a person developing a condition. Especially for dementia, there are a lot of risk factors – some of them are under person’s control (for example smoking, exercise etc.) and some of them are not (for example genes of a person, family history, age etc.). One cannot say that persons having any of the risk factors will necessarily develop dementia, in the same way avoiding any of the factors may not help the person stay healthy, but there are more chances of being healthy by avoiding certain risk factors. As discussed in chapter 1.4, this study chooses a reductionist approach as risk factors related to dementia can be categorised into three different categories. 1) Genetic risk factors 2) Medical factors 3) Non-medical risk factors. Category three is further divided into two categories.
Genetic risk factors related to dementia may include the APOE*E4 which has been associated with increase in risk of late-onset Alzheimer’s, and the three genes Amyloid Precursor Protein (APP), Presenilin-1 (PS-1) and Presenilin-2 (PS-2) which are associated with early onset Alzheimer’s, medical risk factors may include blood pressure, body mass index (BMI), cholesterol, diabetes, hypertension, etc. Non-medical risk factors are the ones which are external to the person’s body, only indirectly affecting health. This may include age, sex, educational years, alcohol consumption, smoking habit, etc.

Genetics is an important dementia risk factor which is not under person’s control. How the information related to genes is helping in treatments and prevention of dementia and specific genes by including in ICT tools is discussed in the following chapter (Chapter 3). There are a lot of medical, non-medical factors that effects dementia, but finding all those factors is out of this project scope. Instead, some of the non-medical factors have been identified (Chapter 5) and discussion of these factors is presented.
3 Targeting the individual

Most of the mobile apps like diet-related apps, risk prediction apps, etc. are targeted towards an individual, so it is important to see that individual as a part of wider development. Tools developed earlier did not give much importance to data related to individuals as most apps were developed from statistical and population-based studies. New technologies developed by ICT helps in improving the quality of life and reducing the effects of ageing [26]. Combinations of genetics, genomics, technology, and therapeutic measures can bring improvements to the healthcare system by targeting individual. Precision medicine is another approach which is also targeting the individual. Genetics is the study of heredity. Genomics is the study related to genes and their functions, and techniques related to them. Genetics is about the functioning and composition of a single gene whereas genomics is about all genes and their relationships in order to identify their combined influence on the development of the organism [27]. The coordination and combination of bioinformatics is helpful in developing various tools for the people who are suffering from dementia. Prevention is possible when we know the risk of people to get diseases. Genomic research is becoming widespread by providing information related to risks of getting specific diseases in early stages by studying molecular aspects of the disease. For instance, studying amyloid-beta protein, the apolipoprotein E (APOE) E4 allele (APOE*E4) helps in determining Alzheimer’s disease risk [28].

3.1 Personal genomics

The full genome is the complete DNA sequence containing the complete information of an individual. An allele is one of the several forms of a gene. Different alleles can result in e.g. different pigmentation. For example, the gene for eye colour has several alleles such as an allele for brown eyes and an allele for blue eyes. Genomics is the study of genes and their functions. The main aim of genomics is understanding the structure of the genome, including the mapping genes and sequencing the DNA. Personal genomics is a branch of genomics which deals with the analysis of the genome of individuals. There are various new techniques available for identifying genotypes of individuals, like full genome sequencing or single-nucleotide polymorphism (SNP) genotyping and identified genotypes are compared with existing literature in order to predict the risk of genetic disorders. Every genome is unique and new sequencing technologies made available to get the sequenced genomes for individuals with fewer costs. When the Human Genome Project started it took seven years to sequence the first percent of the human genome, while the second per cent took only one year. Techno-optimists said that genome sequencing becoming cheap enough, it can be used on every person, every cancer cell, every bacterium, and every virus, which helps in predicting the future [29]. Personal genomics is very helpful for identifying genetic inclination of a particular person towards common diseases. If a person or a caregiver or an insurance company knows that one is inclined to a certain disease it will be easier for them to take precautions such as change lifestyle, diet, physical activities and regular health check-ups [30]. The changes in one of these lifestyle-related factors may influence the total risk score.

Personal genomics can be used by health care to educate and counsel the citizens. It can also be used to treat certain diseases with just the right medication, using existing data about how the substance worked on other patients with same genotypes to ensure that the medicine has maximum effect on the disease with no or fewer side effects. Personal genomics can also be used to advise couples who are planning to have children. For example, if both individuals are carriers of a genetic disorder like cystic fibrosis it will increase the probability of the child having
this disease. This may give a choice to the couple having babies in another way, for example using IVF where embryos are diagnosed in the lab for genetic disorders before implanting in the womb [30].

Mobile applications have been developed to analyse an individual’s DNA by submitting their genome sequence. An example of such an app is GeneG, in which a user can upload their genome sequence to a site and get to know if they have any predispositions for certain diseases by just clicking on a smartphone screen. For example, a pregnant woman needs to donate a sample of DNA for each disease she wants to test since traditional clinics generally process only one gene at a time. This means that the pregnant woman has to go to the clinic every time when she needs to donate sample and takes several days to get the results. Using GeneG user can upload their genome in VCF format (a format used for genome sequence variations) to the site. Once online, they can submit the genetic information to standard analyses developed by organisations like the National Institute of Health, European Bioinformatics, and Standard University. All this can be done remotely by just clicking the app. The information submitted by GeneG user will be kept securely and is under the control of user including who can access the test results. Using information got by GeneG’s testing system doctors can provide successful treatments for various diseases which range from cancer to sleep disorders. At the same time user have more information about their genetic data, they may demand solution according to the results from DNA tests which puts pressure on medical people. Gentle is an iPhone app which offers a DNA-test that includes analysis from 1700 genetic conditions. Many genetic testing companies’ takes the test results and dump into user’s web account without any further counselling. Apart from this Gentle app provides test results information to doctor appointed by the user and further Gentle team provides service for follow-up questioner that may come from user’s doctor. The test data will be transferred securely and user’s data will be kept secretly by only allowing user’s prescribed doctor and geneticists.

The cost of personal genomics is decreasing and dropped to around $1200 in 2015. In Figure 3, the cost curve shows a decrease in genome cost. In 2001 the cost was $95 263 and this was reduced to $1245 by 2015 [31]. Many companies like 23andMe, Navigenics, Pathway genomics, and deCODEme then offered services for genetic testing. Reduction in cost of personal genomics indicates that in near future common people may get their genome sequenced which is helpful in predicting risk related to certain diseases. Risk prediction applications may collect genome information from the user for efficient prediction results.
3.2 Personalised health

Not every disease affects every person in the same way, and even the treatment needs to be varied [32]. In the course of unprecedented scientific innovations and technological enhancements, precision medicine has the ability to detect the disease at its earliest stage (in some cases even before a person is born and still in the stage of the embryo or even before the fertilised egg is implanted on uterus wall) and stop the progression of the disease. The genotyping of drug metabolising enzymes has also helped in improving the drug dosage and thereby reducing the side effects drastically [32 page 4].

Precision medicine uses the molecular markers that signal the symptoms of a disease very earlier than the disease is diagnosed in a clinic. President Obama paid out $215 million in 2015 as an initiative for discovering genetic causes of diseases [33]. As a part of precision medicine initiative President Obama and other officials announced that the National Institutes of Health going to spend $55 million in a single year to find out the genetic and environmental risk factors interaction in causing cancer disease [34]. Many diseases such as dangerous cancers can be
totally cured when diagnosed at early stages. Personalised health not only detects the diseases at the very early stage but also shifts the emphasis of medicine from reaction to prevention.

<table>
<thead>
<tr>
<th>Class</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-depressants (SSRIs)</td>
<td>38%</td>
</tr>
<tr>
<td>Asthma drugs</td>
<td>40%</td>
</tr>
<tr>
<td>Diabetes drugs</td>
<td>43%</td>
</tr>
<tr>
<td>Arthritis drugs</td>
<td>50%</td>
</tr>
<tr>
<td>Alzheimer's drugs</td>
<td>70%</td>
</tr>
<tr>
<td>Cancer drugs</td>
<td>75%</td>
</tr>
</tbody>
</table>

*Figure 4 Percentage of patient population where a particular drug in a class is ineffective [32]*

Personalised health helps avoid adverse drug reactions. Around 5.3 percent of all hospitalisations are due to drug reactions [35]. Precision medicine has fewer side effects and thereby increased patient adherence to treatment. For example, a study revealed that patients with knowledge of a genetic predisposition for high cholesterol have shown more than 86 percent adherence to their treatment, whereas it was only 36 percent of patients without this knowledge [36]. Precision medicine will also improve the quality of life, for example usually a heart transplant patient needs continuous diagnosis for if immune system is trying to reject the new organ (which can be fatal for the patient) and this diagnosis is very complicated where a pipe is inserted into a vein in neck and threading it towards heart. This entire process simply can be replaced with a simple molecular test which requires a blood sample, making life easier and smoother. Personalised health can help in finding alternative uses for medicines. A medicine which is ineffective on a generalised group of people can be very effective in a specific group. So if a medicine which is ineffective in 90% of cases can still be manufactured for 10% of the population and need not disappear from the production. Precision medicine can help to control the overall costs of health care, for example by avoiding trial and error medicine prescription, by reducing the hospitalisations due to adverse drug reactions and by detecting the diseases very early stage (and saving the cost of drugs).

As every coin has two sides, precision medicine has its own limitations. Some difficulties are expressed by physicians due to uncertainty in genomic test results. There are difficulties in putting precision medicine into clinical practice as participants do not have proper training regarding genomic tests and interpretation of test results. Consideration of only genetic risk factors may not be helpful as dementia can be the effect of both genetic and non-genetic risk factors (ex: age, education etc.). One has to think about risks of sharing personal genomic data before using it.
Study related to personal genomics, precision medicine is important as particular allele (example \textit{APOE*E4}) associated with risk of developing dementia. Consideration of personal genomics information in MULTI-MODE app may provide better results which are applicable for individuals.
4 Risk scores

Various risk scores have been developed to predict the risk of individuals for cognitive decline within a given time frame. These risk scores include only a few well-known risk factors that are easily measurable to calculate the subsequent risk of the event or disease. The main use of these risk scores is targeting of the preventive measures to those who are at risk and are also used to distribute understandable information to the layperson. The developed risk scores have been calculated based on different risk factors. For understanding various risk factors related to dementia, some risk scores have been discussed in this section.

Finish Geriatric Intervention Study (FINGER), CAIDE risk scores were described which are the base for new CAIDE risk score for late-life. Old CIADE risk score app has experimented with selected person’s data and results are shown in section 4.2. ANU-ADRI study has been explained, as study involved with some experiments related to old CAIDE risk score. To explore more about risk factors used in late-life dementia, further desktop research has been conducted and lead to two risk scores LOAD and Late-life dementia risk index, as these are related to late-onset dementia and study related to these is easily available on the internet.

4.1 FINGER

FINGER was a two-year population-based multi-domain randomised controlled trial [37] that tests whether nutritional guidance, mental training, exercise, and reduction of vascular risk factors (for example control of blood sugar levels in people with diabetes) can prevent cognitive impairment and disability.

One cannot change one’s age or family histories but there are some modifiable factors which help in active ageing. A third of Alzheimer’s disease is correlated with some modifiable risk factors such as low education, mid-life hypertension, mid-life obesity, diabetes, physical inactivity, smoking, and depression [38]. Better lifestyle, e.g. due to improved economic status can have a positive impact on dementia in later stages. Physical activity, cognitive training, or both have positive effects on cognition when compared to single-domain prevention trail (testing the efficacy of diet, exercise, cognitive stimulation, and vascular risk factor interventions separately) for cognitive impairment [39].

The FINGER study was the first intervention trial in the world using multi-domain intervention approach for preventing dementia. In this trial, participants aged 60-77 years were selected from previous national surveys. Participants were having, CAIDE (Cardiovascular Risk Factors, Ageing and Dementia) Dementia Risk Score of 6 points or more and cognition at the mean level or less than expected for age. Participants were divided into two groups (1:1 ratio), one group is a multi-domain intervention (diet, exercise, cognitive training, vascular risk monitoring) and another group is control group (regular health advise). A two-year study is done and cognition is measured through neuropsychological test battery (NTB) Z-score. Study is done by providing regular health advice to control group and additional four intervention components related to diet (nutritional intervention), exercise (programs for progressive muscle strength training, aerobic exercise), cognitive training (memory tests, word games), vascular risk monitoring (by checking and advising about blood pressure, weight and BMI (Body Mass Index) to intervention group. Various tests have been done in the two-year study at intervals of 6 months, 12 months and 24 months. After 2-year study results are observed as NTB total score (a measure of drug efficacy used in clinical trials), executive functioning (like planning, problem-solving, attention, decision-making), processing speed, memory are in the much better condition in the
intervention group compared to control group. So this proof of concept randomised control trial (RCT) showed the importance of multi-domain prevention approaches which also shows beneficial effects of intensity, long duration, type (e.g., multi-domain), and choice of participants (risk individuals) on cognition. FINGER study shows how long multi-domain intervention has its positive effect in preventing or maintaining cognitive decline but does not explain the effect of the individual factor on the overall effect. Considering the effect of individual factor may give beneficial results in developing evidence-based e-health tools which help in promoting healthy brain in ageing.

### 4.2 CAIDE risk score

The CAIDE risk score was developed to identify the risk level of a person getting dementia in his/her old age. Data for calculating risk score is taken from CAIDE study which is conducted by considering 1409 individuals (875 (62%) were men and 534 (38%) were women from four separate and independent population-based random samples) by examining their mid-life and conducting re-examination 20 years later at late life for later signs of dementia. This study is based on multifactor analysis related to education, age, sex, hypertension, hyperlipidaemia, and obesity [40]. These factors are divided into three categories: vascular (BMI, SBP, DBP, cholesterol), socio-demographic (age, education, sex), and lifestyle characteristics (smoking, physical inactivity). 4% of participants were diagnosed with dementia in those most of the participants were older, less educated, and had vascular risk factors at mid-life compared to the participants who were not diagnosed with dementia [16].

The first survey conducted by CAIDE included basic questionnaire related to health status, health behaviour and medical history. Serum cholesterol, systolic (the top number of blood pressure) and diastolic blood pressure (the bottom number of blood pressure) was taken. Body mass index was calculated from measured height and weight. The second survey (Re-examination) was conducted in 1998 after 20 years from the first survey. During the second survey same survey methods adhered from the previous survey and additionally APOE genotypes were analysed. Cognitive status also considered during the second survey.

Risk scores are calculated from β coefficients using logistic regression models, according to the risk factor profiles at middle age. The range of possible risk scores is 0-15 in model 1 (baseline survey) 0-18 in model 2 (2nd survey). The range is calculated as the sum of a highest possible risk score for each risk factor.

*Table 1 CAIDE Risk Score*

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Risk Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0</td>
</tr>
<tr>
<td>&lt;47 years</td>
<td>0</td>
</tr>
<tr>
<td>47-53 years</td>
<td>3</td>
</tr>
<tr>
<td>&gt;53 years</td>
<td>4</td>
</tr>
<tr>
<td>Education</td>
<td>0</td>
</tr>
<tr>
<td>&gt;= 10 years</td>
<td>0</td>
</tr>
<tr>
<td>7-9 years</td>
<td>2</td>
</tr>
<tr>
<td>0-6 years</td>
<td>3</td>
</tr>
<tr>
<td>Sex</td>
<td>0</td>
</tr>
<tr>
<td>Female</td>
<td>0</td>
</tr>
<tr>
<td>Male</td>
<td>1</td>
</tr>
<tr>
<td>Systolic Blood Pressure</td>
<td>0</td>
</tr>
<tr>
<td>&lt;= 140 mm Hg</td>
<td>0</td>
</tr>
</tbody>
</table>
Example: Consider a 50-year-old woman who is having 8 years of education, physically active, having systolic blood pressure of >140 mm Hg, body mass index of >30 kg/m2, cholesterol of >6.5 mmol/L. Risk score for this person is the sum of the highest possible risk score i.e. 11 (3 + 2 + 0 + 2 + 2 + 2 + 0).

Numerous apps developed using mobile health (mHealth) technologies developing numerous apps for depression, diabetes, cardiovascular diseases and lifestyle changes [41] [42]. To make CAIDE Risk score more accessible to general people and health care practitioners, the researchers who developed CAIDE Risk Score sought to develop a mobile application using mHealth technologies. The result of this is CAIDE Risk Score mobile app which is the first evidence-based App for dementia. CAIDE Risk Score App helps users to detect their individual risk of getting dementia in later 20 years, provides guidance for reducing the risk of getting dementia and also advice for consulting health care practitioner if needed. mHealth apps are helpful in various ways such as keeping track of individual data frequently which helps more data available for health care system for proper medication, tracked data can be used for research purpose in health care systems. One has to consider limitations regarding mHealth apps before releasing into the market. Collected health data from users must be filtered and translated to a message so the user and clinicians understand it. Clinicians and user should be educated before the app being used, to avoid a chance of misunderstanding about the prediction results. Costs involved in developing mHealth tools are also considered as one of the limitation.

CAIDE Risk Score App was released in two versions, one for health care personnel and other for general people [16]. The app is downloaded based on the user selection. Health care practitioner version has an option to enter his/her name, contact details of practice and website information. The app can be downloaded from App Store on iPhone or iPad. Users of App are asked to enter the date of birth, sex, duration of education in years (Figure 5), height, weight, systolic and diastolic blood pressure, serum cholesterol, and physical activity information (Figure 5).
Based on entered information, App calculates the risk score and displays the risk level in the form of a graph. Graph display will be in various colours based on the risk score. For normal risk score (of 8-9) an orange bar, lower risk score (of 0-9) green bar, higher than average risk score (10-15) a red bar (Figure 6) appears.

One of the tabs in the app has the explanation of risk score and provides probability percentage of developing dementia compare to average probability of developing dementia (Figure 7). App also provides the information on how to modify the risk factors so that individual can reduce the chance of developing dementia (Figure 7).
4.3 Australian National University AD Risk Index (ANU-ADRI)

ANU-ADRI Risk Index was developed to identify the degree of older individuals who are at risk of AD and an attempt made to compare the risk index with old CAIDE risk score which was for middle-aged people [17]. This tool differs from previous tools developed to predict dementia, as it was developed by identifying risk factors from three independent older cohorts instead of risk factors from a single cohort study, and did not include variables which require laboratory tests. The tool has wider utility as it was evaluated against both subtypes of dementia like AD and general dementia. An evidence-based medicine approach was selected to identify the risk factors (age, sex, low education, traumatic brain injury, diabetes, low social networks, depressive symptoms, smoking) and protective factors (cognitively stimulating activities, physical activity, alcohol consumption, fish intake) for AD. Developed risk index excluded BMI and. ANU-ADRI risk score had a range of -11 to 56, which involved complete data from 2496 participants. In a study made by ANU-ADRI, CAIDE risk score was calculated and compared in three different older aged cohorts. The comparison of CAIDE risk score did not result in high c-statistics when used on older cohorts. This can be due to CAIDE development was done by considering mid-life risk factors and was validated in older cohorts. The effect of risk factors can be different in different ages. Factors like BMI, and cholesterol less associated with late-life AD risk [43] [44] when compared to mid-life and were excluded from ANU-ADRI.
### 4.4 Comparison between old CAIDE, ANU-ADRI and new CAIDE risk scores:

Table 2 Comparison of risk scores

<table>
<thead>
<tr>
<th></th>
<th>Old CAIDE risk score</th>
<th>ANU-ADRI risk index</th>
<th>New CAIDE risk score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Based on mid-life risk factors</td>
<td>Based on late-life risk factors</td>
<td>Based on late-life risk factors</td>
</tr>
<tr>
<td></td>
<td>From single cohort study (CAIDE)</td>
<td>From three independent cohorts of older adults (MAP, KP, CVHS)</td>
<td>From three independent cohorts of older adults (SNAC-K, CAIDE, KP)</td>
</tr>
<tr>
<td></td>
<td>Score values were derived from β coefficients of the logistic regression model.</td>
<td>Score values were derived from β coefficients of the logistic regression model.</td>
<td>Score values were derived from β coefficients of cox regression model.</td>
</tr>
<tr>
<td></td>
<td>1409 participants</td>
<td>2496 participants</td>
<td>3744 participants</td>
</tr>
<tr>
<td></td>
<td>Risk score range is 0 to 15</td>
<td>Risk score range is -11 to 56</td>
<td>Risk score range is 0-15</td>
</tr>
<tr>
<td></td>
<td>Preliminary study: Age 39-64 yrs. Follow-up study: 65-80 yrs.</td>
<td>Comparison of adults aged &lt;70 with &gt;=70 was done</td>
<td>Comparison of adults &lt;74 with &gt;= 75 was planned</td>
</tr>
</tbody>
</table>

**Strengths:**

1. Information gathering done at mid-life and again at late-life
2. Number of participants was high: more than 80% of the baseline examination and greater than 70% at the follow-up examination.

**Limitations:**

1. The score is applicable to the prediction of dementia risk for the people who survive 20 years after the baseline study. Those who died before follow-up study
2. ANU-ADRI did not validate against younger cohorts.
3. Isolated findings from KP, MAP, and CVHS cohorts were used in some of the meta-analyses but validation

**Strengths:**

1. Data collected from three independent older cohorts.
2. The tool has wider utility as evaluated against subtypes of dementia like AD and general dementia.
3. Risk factors do not need any laboratory tests.
4. Risk index has been compared with pre-developed CAIDE risk score.

**Limitations:**

1. ANU-ADRI did not validate against younger cohorts.
2. New risk score has been validated according to TRIPOD rules.
may have had dementia which was not considered.

2. The family history of dementia, waist-hip ratio and presence of diabetes or insulin resistance, concentrations of high-density and low-density lipoproteins were not considered in calculating risk score which can be a part of developing dementia.  

samples were not purely independent.

<table>
<thead>
<tr>
<th>MAP - Memory and Aging Project, KP - Kungsholmen Project, CVHS - Cardiovascular Health Cognition Study</th>
</tr>
</thead>
</table>

### 4.5 Late Onset Alzheimer’s disease risk score

A risk score has been developed for the prediction of Alzheimer’s disease in elderly persons based on vascular risk profiles [45]. Participants for the study were the people aged 65 years or older and were residents of New York who were free of dementia and cognitive decline. Age, sex, education, ethnicity, APOE*E4, diabetes history, high-density lipoprotein levels, hypertension or smoking and waist to hip ratio were considered as risk factors. Several risk factors were explored separately with Late Onset Alzheimer’s disease risk score (LOAD) using Cox proportional hazards models to identify the risk factors which contributes more to the risk score.

All participants in the study underwent an interview related to general health, medical history, and neuropsychological issues. 1051 people participated in the study. For calculating risk score, study chose the similar approach used in CAIDE. A follow-up study conducted after 18 months of base study. Follow-up study resulted in less number of people with dementia compared to the first study. This is due to the demented people from the first study were older, less educated, had a higher prevalence of diabetes, a higher WHR (waist to hip ratio), and lower HDL-C (high-density lipoprotein cholesterol) levels.

Risk score has been categorised into five groups due to large risk score range (0-60). According to developed risk score quintiles, the probability of LOAD was 1.0 persons with a score of 0-14, 3.7 for persons with a score of 15-18, 3.6 for persons with a score of 19-22, 12.6 for persons with a score of 23-28, and 20.5 persons with a score higher than 28. The probability of LOAD was increased with higher vascular risk score and a greater number of risk factors.
### Table 3 LOAD Risk score

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Risk Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
</tr>
<tr>
<td>Age (yrs.)</td>
<td></td>
</tr>
<tr>
<td>65-70</td>
<td>0</td>
</tr>
<tr>
<td>&gt;70-75</td>
<td>6</td>
</tr>
<tr>
<td>&gt;75-80</td>
<td>8</td>
</tr>
<tr>
<td>&gt;80-85</td>
<td>13</td>
</tr>
<tr>
<td>&gt;85</td>
<td>21</td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Current Smoking</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td>Low HDL-C</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>High WHR</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Yes</td>
<td>7</td>
</tr>
<tr>
<td>Education (yrs.)</td>
<td></td>
</tr>
<tr>
<td>&gt;9</td>
<td>0</td>
</tr>
<tr>
<td>7-9</td>
<td>8</td>
</tr>
<tr>
<td>0-6</td>
<td>11</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0</td>
</tr>
<tr>
<td>Black</td>
<td>5</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4</td>
</tr>
<tr>
<td>APOE*E4</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0</td>
</tr>
<tr>
<td>&gt;= 1</td>
<td>4</td>
</tr>
</tbody>
</table>

This risk score is mainly used to estimate the risk of developing dementia in older people and it can also use in genetic research of dementia to adjust a compound variable of non-genetic risk factors.

### 4.6 Late-life dementia risk index

The late-life dementia risk index can accurately divide older adults into those with a low, moderate, or high risk of developing dementia within 6 years [46]. 3,375 participants were involved in the Cardiovascular Health Cognition study [47] without evidence of dementia at baseline. Logistic regression models are used to identify the risk factors which are most predictive of developing dementia within 6 years and developed a point base system. The
point’s possible range is 0-15. Participants mean age baseline was 76 years; 59% were woman and 15% were African American. Fourteen percent (n=480) of the participants developed dementia within 6 years. The developed point based system includes the risk factors age, poor cognitive test performance, body mass index, APOE*E4, cerebral MRI findings of white matter disease, internal carotid artery thickening on ultrasound, slow physical performance, history of bypass surgery, lack of alcohol consumption. 4% of the participants subjected with low scores, 23% of participants subjected with moderate scores and 56% of participants subjected with high scores.

Table 4 Late-life dementia risk score

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>75-79 (yrs.)</td>
<td>1</td>
</tr>
<tr>
<td>80-100 (yrs.)</td>
<td>2</td>
</tr>
<tr>
<td>Low 3MS</td>
<td>2</td>
</tr>
<tr>
<td>Low DSST</td>
<td>2</td>
</tr>
<tr>
<td>BMI &lt; 18.5</td>
<td>2</td>
</tr>
<tr>
<td>&gt;=1 APOE*E4</td>
<td>1</td>
</tr>
<tr>
<td>MRI white matter disease (grade&gt;=3)</td>
<td>1</td>
</tr>
<tr>
<td>MRI enlarged ventricles (grade &gt;=4)</td>
<td>1</td>
</tr>
<tr>
<td>Internal carotid artery thickening &gt;=2.2mm</td>
<td>1</td>
</tr>
<tr>
<td>History of coronary bypass surgery</td>
<td>1</td>
</tr>
<tr>
<td>Time to put on and button shirt &gt;45s</td>
<td>1</td>
</tr>
<tr>
<td>Lack of alcohol consumption</td>
<td>1</td>
</tr>
</tbody>
</table>

Age is compared to those who aged 65 to 74 years.

Low 3MS: <= 87 (all white participants and black participants with >=12 years education) or <=70 (black participants with <12 years of education).

Low DSST: <=33 (white participants with >=12 years of education) or <=22 (white participants with <12 years of education and all black participants).

3MS: Modified Mini-Mental State Examination; DSST: Digit Symbol Substitution; BMI: Body Mass Index; CI: Confidence interval.

The late-life dementia risk index tool can be used in clinical or research area to target prevention and intervention strategies towards high-risk individuals. The late-life dementia risk index also used to identify the older adults who should be monitored for the new dementia symptoms, so that treatments could be initiated at the earliest possible stage of the diseases, and helps in providing information to concerned patients or family members. The late-life dementia risk index could be used to reassure the older adults who do not currently have overt dementia and to provide those individuals whose risk is high with information that may helpful in planning for better future.

4.7 MULTI-MODE

To promote healthy living, improve health care, and support active ageing, EIT (a part of the European Union Horizon 2020 effort) has launched EIT Health. EIT Health integrates higher education, research, and business for the enhancement of healthy living and to improve health
care in Europe. Every country in Europe has different healthcare systems. EIT Health is concentrating on developing products or healthcare solutions that help in the integration of healthcare provision. EIT Health centres are located in different regions of Europe which helps for better results of fragmented healthcare systems. By the end of 2018, EIT Health aims to create 165 start-ups and support 160 new service and product solutions. As part of this, EIT Health launched the MULTI-MODE project in which researchers at SICS are planning to develop an evidence-based tool (new CAIDE tool) for improving and maintaining healthy brain in ageing and will be validated according to European standards. This tool is going to be used by citizens and healthcare providers to decrease healthcare costs and social burden. Existing risk scores have been calculated using medical data. There are some non-medical factors which may contribute to the risk score calculation. Identification of such non-medical factors and including these non-medical factors in further risk scores may improve the predictability of developing risk score.

CAIDE Risk Score App (described in Chapter 4.2) is the first evidence-based tool to predict the dementia risk in individuals at their late life, having been validated on a U.S. multi-ethnic population [48]. New mid-life risk factors like obesity, depression, diabetes mellitus, smoking, poor lung functioning, and head trauma were added to CAIDE Risk Score to increase predictability. Risk model prediction and validation were examined using the C-statistic, a measurement of the probability that predicting the outcome is better than chance and is used to compare the goodness of fit of logistic regression models which measures in between 0.5 to 1, where 0.5 means the model is no better than chance and 1 means model identifies the factors correctly. It also tests net reclassification improvement (a measure of correctness of upward and downward movement as a result of adding new factor) and integrated discrimination improvement (is used to estimate the change offered by the addition of a new factor to prediction model, e.g. dementia, after introducing a new factor into the model). The experiment resulted in a C-statistic of 0.75, which is similar to the original CAIDE C-statistic of 0.78. The result of adding new risk factors did not improve its predictability [48]. This may be due to not considering correlated risk factors properly: there may be a change in output when risk factor combinations differ. On the other hand, there are studies conducted on mid-life diabetes that concluded that diabetes at mid-life has a stronger risk of dementia when compared to diabetes at late-life [49], so there can be a change in risk score considering diabetes. CAIDE study includes the risk factors related to dementia and AD. CAIDE Risk Score App was designed to predict dementia particularly AD, most of the risk factors included in the app includes risk factors related to dementia in general but not particular to AD.

As explained in section 4.3, analysis of CAIDE risk score in older cohorts showed poor results. This reveals that there is a need for a new risk score which fits better to an older population. There is an opportunity to consider factors related to non-medical data while calculating the new risk score. From studying the population from Kungsholmen, Stockholm [SNAC-K] and considering some additional factors to CAIDE, a new composite risk score has been developed by researchers at KI/ARC which can be suitable for older cohorts. To judge the correctness of the new developed CAIDE risk score, the developed risk score has been validated according to “Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis” (TRIPOD). The TRIPOD statement aims to improve the transparency of the reporting of a prediction model (developed to help healthcare staff in estimating the risk of disease) study regardless of the study methods used. The initiative developed a set of recommendations that have to be followed by reporting of studies when developing, validating, or updating prediction model. A great number of recommendations have been listed after a Web-based survey and
revised by journal editors and health care professionals. These recommendations have been reduced to 22 after several meetings and e-mail discussions. TRIPOD is becoming widespread as there is a need for recommendations which helps to judge the quality of reporting of prediction studies. The TRIPOD statement has been published in several journals like BJOG, British Journal of Cancer, British Journal of Surgery, BMC Medicine, British Medical Journal, Circulation, Diabetic Medicine, European Urology, European Journal of Clinical Investigation, and Journal of Clinical Epidemiology [50]. Researchers from SICS are going to develop a mobile app based on the new risk score. MULTI-MODE is a project which is planned for three years. During the first year, the researchers have developed a new risk score based on the study results which is suitable for the older population, meanwhile, researchers at SICS started developing the preliminary version of the risk score app (i.e. 2015-2016). The second year is planned for validating developed risk score and fully working app version based on developed risk score (i.e. 2016-2017).
5 Non-medical risk factors

Adding extra risk factors to an existing risk score like CAIDE can improve the performance of predictability tools. This chapter is going to discuss additional non-medical risk factors that can be considered while calculating risk scores. Answering the following questions related to non-medical risk factors can be helpful to understand the importance of these factors in calculating risk scores.

1. **What is non-medical data?**

   Non-medical risk factors are the ones which are outside the person’s body (outside the area of medical factors) but which affect the health status such as environmental factors and lifestyle related factors. Examples can be educational years, smoking, alcohol consumption, sleep deprivation, etc. that affect a person’s health.

2. **How does non-health data potentially contribute to computing correct individual risk scores?**

   There are many non-medical risk factors which increase the risk of dementia. Depression can be considered as an example, which plays a role in developing dementia [51]. Physical activity also plays a role in developing dementia [52]. Maintenance of physical fitness and good eating habits [53] helps in lowering the risk of dementia. Not considering the effect of such non-medical risk factors may produce an inaccurate or incomplete risk score.

3. **Why has so little previous research and development taken non-health data into account?**

   This may be due to researchers being unaware of the effect of these non-health data and that this data was not widely available before, whereas medical data was available easily from electronic health records (EHRs). Nowadays, many personal devices are available to store a person’s activities data which can be helpful in tracking and collecting data related to non-medical factors. More amounts of data can be collected from apps when compared to medical health records, as apps can collect the data at short intervals of time. Considering mental-health data, this is hard to get from medical health records. Apps are helpful in collecting such information. Many mobile apps are releasing into the market related to mental-health. If we type ‘depression’ into the Apple App Store then a list of at least hundred apps pop up on the screen. There are apps that diagnose depression (Depression test), track moods (Optimism), etc. There is much data related to depression available from these apps. These mobile apps collect the data from users regularly and they also interact with users proactively and pinging them to ask about their moods, thoughts, and overall well-being. Several times a day, the app might ask questions such as “How are you feeling today?”, “How well did you sleep last night?”

Combining non-medical data with medical data makes models more complicated. To keep the models simple, some of the non-medical factors which are hard to measure are avoided by researchers. Considering ‘social isolation at old age’ risk factor, which was a bit complicated to
measure in olden days. Services like befriending, PatientsLikeMe (website for finding similar patients) etc. are helpful in measuring these factors. Befrienders are volunteers who are interested in meeting new people, supporting and helping them with good conversation, the occasional trip out, or providing an opportunity to participating in social activities. There are two types of befriending, one is telephone befriending and other is visiting befriending. Now with help of befriender’s service one can measure the social activities related to the particular person. Like how many new people a person meeting per day? / Per month? How many times per day he/she getting calls from befriender? Measures related to these factors helps in introducing this non-medical factor into dementia risk calculator.

Costs related to collecting, storing, and analysing data decreasing dramatically. On the other hand, there are large amounts of data being generated. Genomics and usage of health apps producing a large amount of data, this opens the door for integration of dementia with big data. Big data can be helpful in analysing and producing predictive tools with the consideration of non-medical risk factors.

The non-medical risk factors can be divided mainly into two categories; risk factors related to the person for whom the risk score is calculated, and app- and user-related (technical) risk factors. The first category can be further divided into geographical risk factors, psychological risk factors, and lifestyle-related risk factors. WHITE PAPER [8] was helpful in categorisation of these risk factors.

5.1 Geographical related risk factors

5.1.1 Air pollution
A study conducted on mice which have been exposed to air pollution specifically to a nickel nanoparticle, a component of air, exposed mice rapidly developed Alzheimer’s with doubling its speed. Air pollution causes changes to hippocampus where learning and memory processing takes place. Possibly due to city’s air pollution, children of Mexico have developed early signs of Alzheimer’s disease [54]. The air pollution need not be outside one. Most of the Alzheimer’s patients in China (whose is also the world’s largest consumer of tobacco) were exposed to the second-hand cigarette smoke [55]. A recent study tried to combine medical factor with a non-medical factor like air pollution. [56] Investigated association between the medical precision of drugs to young people and thus correlated with air pollution which is a non-medical one. The study concluded that there could be a link between exposure to air pollution and medications used to treat the psychological disorders in children.

5.1.2 Geography (Area of living/Ethnicity)
Environmental and cultural factors may have a significant effect on Alzheimer’s disease. The most common diseases in Western Europe and the United States is Alzheimer’s disease, while vascular dementia is the common one in East Asia. Consideration of high-temperature exposure, how many hours the person exposed to temperatures over 40ºC or how many hours person got exposed to harmful radiation may have an effect on demented people’s death [57].

5.2 Psychological risk factors

5.2.1 Developmental and early-life risk factors / Stress level
Negative experiences in early life can cause stress, which could affect the brain structure. There can be an increased risk of dementia or AD for those who experienced an early parental death, for instance [58]. A prospective study has been conducted in Sweden to associate the stress caused in early-life with late-life dementia. 374 people who are aged 70 and over, residents of
Göteborg participated and the follow-up was carried out for 9 years. Data related to five stress factors (early parental death, growing with one parent, divorce of the parents, growth under different guardians, poverty) were collected. Data related to people who faced any of these five mentioned factors was compared with the people who did not face any. The study revealed that the participants who had lost a parent before the age of 16 had a greater chance of developing dementia in late-life (after 70 years) [59].

Physically damaged brains can be one of the causes of stress. Climate changes also have a great impact on stress. There is a chance of stress more on the citizens of both hot countries and cold countries. Rising in temperature forms more stress on people who are already psychologically fragile. Same with colder countries, people spends unproductive time inside due to very cold weather instead of doing activities like hunting, fishing which leads to more stress on elderly people. On the other hand, cold countries have activities like skiing, skating etc. Participating in such activities may help in keeping controlled stress levels. Stress levels may vary according to the area of living.

5.2.2 Depression

Depression is a common mental health conditions in adults. Depression increases the risk of dementia [51]. Depression in older adults can be due to disability in self or family member, social isolation, poor financial status, and change in living situation [60]. Sleep disturbances can be a risk factor for depression in older adults [61]. A number of ICT solutions have been produced to reduce the depression on older adults by encouraging social activities, some form of games or achievements.

5.2.3 Sleep deprivation

One of the causes for memory loss is the result of neuronal death caused by the increase of toxic plaques, which are an aggregation of amyloid-beta peptides [62]. Sleep helps in wash away toxic proteins at night, preventing them from building up and from destroying brain cells [63]. Studies also show that toxic metabolites, including amyloid-beta, are flushed out of the brain during sleep [62]. If a person has less sleep then there is a less chance of clearing out these toxic proteins. With the help of powerful brain techniques, neuroscientists found that participants with the greatest levels of amyloid-beta had the poorest quality of sleep and performed worst at memory tests [64]. There is no answer which one is first, poor sleep leads to less chance of clearing out toxic proteins or increase in the amount of toxic proteins leads to poor sleep.

5.3 Lifestyle-related risk factors

There are some lifestyle related factors which affect dementia. Examples of lifestyle-related factors can be how much time person active in hobbies like painting, music, kind of movies person likes, interested games (can be physical games or online games), how much code person uploaded to GitHub or how many queries he/she answered on stack Overflow (A site used by programmers for discussing language related issues), the type of transport person uses (whether person uses bike or collective traffic like buses and trains).

5.3.1 Smoking

It is widely accepted among the researchers that the use of tobacco may increase the risk for Alzheimer’s disease. Smoking can lead to stroke, cerebrovascular disease. Smoking increases the total plasma homocysteine, which is the main risk factor for stroke, Alzheimer’s disease and cognitive impairment [65] [66] [67] [68]. Smoking accelerates atherosclerosis in heart and brain, which prevents usage of brain cells of oxygen and important nutrients [69]. Smoking can cause
oxidative stress, excitotoxicity, neutral death. Oxidative stress leads to inflammation that may directly or indirectly relate to brain changes seen in Alzheimer’s patients [70], [71]. Smoking tobacco increases the APOE*E4 which is a generic risk factor for dementia. Previous studies have evidence that smoking leads to higher risk of developing dementia and Alzheimer’s disease.

5.3.2 Alcohol consumption

Alcohol consumption may increase the risk of dementia and Alzheimer disease. Heavy alcohol consumption can lead to both structural and functional abnormalities of the human brain [72]. There is U- or J-shaped relationship between alcohol consumption and dementia, low to moderate drinking levels reduce the overall risk of dementia whereas high drinking levels increase the risk of dementia [73]. Effect of alcohol drinking on dementia depends on APOE*E4. APOE*E4 carriers, who consumes a light and moderate level of alcohol was associated with greater decline in learning and memory, whereas the light and a moderate level of consumption was associated with an increase in learning and memory among non-APOE*E4 [74]. As discussed in chapter 3, personal genomics is helpful in identifying a person with APOE*E4.

5.3.3 Mid-life coffee and tea drinking and the risk of late-life dementia

Caffeine has stimulating effects on central nervous system in short-term. However, long-term effects of caffeine on cognition are less clear [75]. According to the study [75] which was conducted to know the effect of caffeine on the nervous system in long-term resulted that tea drinking was not associated with dementia. Coffee drinking at mid-life may help in decreasing the risk of dementia at late-life [75].

5.3.4 Social isolation at old age

Most of the loneliness developed at elder age is due to reduced mobility, chronic illnesses and also the loss of spouse or friends. This makes people isolated from social life and increases the loneliness which can be a significant factor in developing a chronic disease like Alzheimer’s. Recent projects suggest technologies for reducing social isolation. The Social Media for All Elderly people (SoMedAll) was designed to make the use of ICT at home for ageing people to deliver happiness, safety and security, education, communication activities [76]. There are some tools developed (ex: JIVE: is a proof of concept for a new communication device) for elderly who are not interested in using complex programs to connect their family and friends. Recent projects investigating robotic solutions for reducing loneliness.

5.3.5 Lack of engagement in reading and hobbies

Higher levels of engagement in reading books and hobbies lower the risk of dementia. Crossword puzzles, baking, painting, gardening, crafts, and playing musical instruments are the examples of leisure activities which keep the brain active and decrease the risk of dementia [77]. According to the study [77], engaging in activities for 1 hour or more daily can protect people from dementia.
6 Analysis

For a better understanding of introducing non-medical risk factors into the risk score, an attempt is made to describe the old CAIDE risk score algorithm.

6.1 Old CAIDE risk score algorithm

![Algorithm for the old CAIDE risk score.](image-url)
Figure 8 illustrates the risk score algorithm based on CAIDE study results. The risk score is calculated in different stages in a step by step manner. In the beginning the risk score is always zero irrespective of who the app user is, in the first and the most important step, the algorithm adds risk score from age, if the age is under 47 then a message will be displayed that app is made to calculate risk score for those who are over 47 years old and then stops. In step2 the score is added from education. In further steps, scores will be added considering Gender, Blood pressure, BMI, Cholesterol levels in the blood, Physical activity and APOE at the end.

6.2 Introducing non-medical risk factors into risk score

This chapter has the analysis of introducing some non-medical risk factors (one risk factor from each risk factor category) into the calculation of risk score, and checking the data entered by the user in improving risk score. Due to unavailability of data related to these factors, some assumptions were made during the study.

Considering air pollution from geographical risk factors:

As described in Chapter 5.1.1, studies have been conducted related to effects of air pollution on brain-related diseases. Here we are trying to introduce this air pollution into the risk score with the following example. Assume that the following data is entered into the app by a person named as Henrik Svensson.

Table 5 Example data entered by a fictive person into Risk App

<table>
<thead>
<tr>
<th>Factor</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>53</td>
</tr>
<tr>
<td>Education</td>
<td>9 years</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>145</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>90</td>
</tr>
<tr>
<td>BMI</td>
<td>40</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>7</td>
</tr>
<tr>
<td>Physical activities</td>
<td>No</td>
</tr>
</tbody>
</table>

Now from the algorithm shown in Figure 8, we can get the risk score as follows

Table 6 Risk scores for fictive data in table 5

<table>
<thead>
<tr>
<th>Factor</th>
<th>Value</th>
<th>Risk score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>53</td>
<td>3</td>
</tr>
<tr>
<td>Education</td>
<td>9 years</td>
<td>3</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>0</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>145</td>
<td>2</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>BMI</td>
<td>40</td>
<td>2</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Physical activities</td>
<td>No</td>
<td>1</td>
</tr>
<tr>
<td>Total Risk Score</td>
<td></td>
<td>12</td>
</tr>
</tbody>
</table>

Some of the non-medical factors such as air pollution are difficult to relate to dementia or Alzheimer’s disease to calculate the risk score, as air pollution is a general term and it does not include which chemical component in the air is higher and which is lower. One possibility can
be co-relation between Air Quality Health Index (AQI) vs risk score, but in this case the study should include participants from around the world as a single country like Sweden cannot have places with all possible air quality indexes (In fact most of the Swedish cities have very good air quality compared to rest of the world) so it would be quite unfair to draw such correlations using a survey from a single country. Risk score app could become a good tool for this kind of survey as the app can be accessible around the world, and if we can the information related to client location at server side and risk score, it is possible to get air quality index for that particular location from some other sources and analyse the risk score to draw any correlation between risk score and Air quality. Consider the following example of Air Quality Index and its effect on the risk score. The Air Quality Index data shown in the following example is from 20th August 2016.

**Table 7 Example data for Air Quality Index**

<table>
<thead>
<tr>
<th>City</th>
<th>Air Quality Index (PM 2.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stockholm, Sweden</td>
<td>10</td>
</tr>
<tr>
<td>Bangalore, India</td>
<td>171</td>
</tr>
<tr>
<td>Beijing, China</td>
<td>168</td>
</tr>
</tbody>
</table>

As discussed in earlier chapters, air pollution may increase the risk of dementia. Here the assumption is made that for every 50 number of AQI risk score increases by 1. From the table 6, the risk score is calculated as 12. The risk score may not change for the person who is living in Stockholm, Sweden as AQI is very low i.e. only 10. But considering the person living in Bangalore, India or Beijing, China who were exposed to higher AQI, the risk score may increase approximately by 3 (171/50 = 3.42, 168/50 = 3.36) and the person's risk score may increase from 12 to 15 (considering integer part). In this way, considering air pollution in the calculation of risk score may affect the person's overall risk for dementia. AQI varies daily. While considering data in risk calculation it is better to consider average AQI data instead of current data.

One practical problem could be that if a person mentions his address, for example, Stockholm, does not mean that he spent whole his life in Stockholm, he could have spent many years in low air quality places and the equations could get very complicated, but then again it is very difficult to collect all of this information and analyse unless someone conducts a detailed survey which contains thousands of questions. The introduction of air pollution risk factor into the risk score calculation is associated with certain costs and efforts, as it needs collection of information from participants from all over the world and the data related to the air quality health index for different cities should either be fetched from other third party data sources or by maintaining it in an own database and updating it frequently (As air quality health index might change often). The calculated risk score must be validated by medical people in order to make sure that the change in risk score is due to the introduction of air pollution and is not influenced by any other factors.

**Considering stress level from psychological risk factors:**

Psychological stress may increase the risk of dementia. This stress can be short term or long term. Stress may come from early life events such as parental death, the divorce of the parents or long-term work related stress. Even though retirement life is supposed to be a peaceful one, there may be other chances of stress, related to changes in work routines, the death of loved ones, poor health condition etc. Considering stress in risk score calculation may change the
overall risk score. Cortisol is a hormone that is associated with stress levels in a human which is used as stress measurement. Cortisol can be measured in blood, saliva or hair. Cortisol measures change frequently, as a function of time of day, therefore interpretation can be difficult. One needs to measure cortisol several times per day to get a better picture of stress than one single measurement. One can see the person’s stress directly by looking at him or her, but making it into a quantitative one is a difficult task. A lot of research has been conducted related to measuring psychological stress, which came up with different solutions. One of the research resulted in measuring stress using a smartphone application from the samples of human saliva [78]. Using this kind of applications, one can measure the stress levels related to a particular user. In future, risk score apps may include one field for stress levels and can connect this field to other apps that can measure stress levels.

**Considering smoking from lifestyle-related risk factors:**

Many studies suggest that smoking has a direct influence on dementia, i.e. higher dementia risk in people who smoke compared with non-smokers. Even passive smoking can increase the risk of dementia. People who smoke have 40 to 80% higher probability of developing dementia than non-smokers[79].

Many of the studies conducted to find out the relation between smoking and dementia are observational studies, meaning N number of participants will be studied for a certain duration (some years) in which they will be regularly checked for dementia. As these studies are observational, it is highly impossible to be 100% sure that dementia is caused directly by smoking or not. Smoking can influence other factors such as cardiovascular diseases, diabetes and some other health conditions which can have a direct influence on dementia.

The smoking can be measured in terms of for example number of cigarettes per day, the number of smoking hours per day or amount of tobacco consumption. There are around 4000 chemicals in a cigarette like Nicotine, Carbon monoxide, Hydrogen cyanide and many other toxins. So it is very difficult to study which of the chemical components are actually influencing dementia at the most. A positive aspect of smoking is that it can be helpful in increasing social network and thereby reduce the stress levels. So while calculating the risk score for smoking, the studies should consider the net score (by considering both positive and negative effects).

If the Risk App can collect the smoking measurements (which are described in the above list) information from its users, and if we can make a big-data using this data collected from the Risk App, this data could be very useful in drawing a direct relation between smoking and dementia. However, Big-data itself can’t be used alone to draw any final conclusions because the analysis results from the Big-data should be always validated in a longitudinal medical study, otherwise the results from the analysis couldn’t be trustworthy.

**Validating non-medical factors (age, sex, years of education)**

The goal of this thesis work was to improve the risk score calculated by CAIDE Risk App by considering factors that were actually not considered in CAIDE study. One of the important non-medical factors which can be categorised as a technical factor that can influence the risk score is erroneous data input into the Risk App. The user of the app can purposely or non-purposely enter incorrect data into the app; while some of the data may not have a big difference on calculated risk score, other data can. According to table 6, calculated risk score is 12.
Now let us assume that the app has an optional text field where the user can enter their personal number if they want and Henrik has entered (or we have collected the personal number from the device with Henrik’s consent) his personal number as 19600101-5657. Now we can calculate the age of the person at client side with a simple javascript function below.

Table 8 Example Javascript function which demonstrates the calculation of user’s age using Swedish personal number or date of birth of the user

```javascript
function calculateAge(personnummer) {
    var birthYear = personnummer.substring(0, 4);
    var birthMonth = personnummer.substring(4, 6) - 1;
    var birthDay = personnummer.substring(6, 8);
    var currentDate = new Date();
    var birthdate = new Date(birthYear, birthMonth, birthDay);
    var age = currentDate.getFullYear() - birthdate.getFullYear();
    var m = currentDate.getMonth() - birthdate.getMonth();
    if (m < 0 || (m === 0 && currentDate.getDate() < birthdate.getDate())) {
        age--;
    }
    return age;
}
```

Before begin with calculating age from the personal number, we can actually validate the personal number by calculating the checksum using Checksum algorithm and comparing it with the last digit (19600101-5657) in personal number. Once the number is validated we can go on calculating Age using the above-mentioned script and the age will be 56. By validating the third digit in the last four digits (19600101-5657, odd numbers for Male and even numbers for Female) we could further see that the user actually entered a wrong Gender, given this personal number.

All of the above-mentioned validations and corrections can be made at the client side without access to any further information. But in order to validate certain things, for example, Education, we cannot do this client-side in a straightforward way without hacking into the client’s device and making some strong assumptions. The server side validation comes into the picture now and it should be a quite straightforward process using the personal number the server will fetch the person’s educational information from BEDA (“Sveriges nationella betygsgesdatabasen”, the Swedish national grades database) and LADOK (Lokalt ADB-baserat studiedokumentationssystem, the local IT-based study documentation system) which is maintained by each university in Sweden. Now assume that we found out that the person is holding a PhD degree from KTH via LADOK, which would mean Education>10 years.

After validation, the risk score from Age will be adjusted to 4, gender-related risk score will be 1, and education-related risk score will be 0. So the adjusted total will be 11 whereas the original risk score was 12. So by using ICT-based validation via non-medical data, we have increased the score, and if we can validate all of the parameters then we can further improve the adequacy of the score.

6.3 Server-side data

The following class diagram shows different objects needed at the server side in order to save the information to the database after the risk score has been calculated at the server side. The objects can either be created at server side or created and serialised at client side before sending it to the server. Figure 9 lists the different fields where some information is sent from the client, and other information that needs to be collected from different server-side sources.
6.4 Database design

This section describes a suggested database design, i.e. tables, fields, relations between different tables, needed to store the information at the server side.

6.4.1 To store client (patient) information

Figure 10 illustrates tables required to store client information: regClient is a table where every client (patient) is assigned a unique id to identify by. Other information could be, name, date of birth, personal number, and address. The address tables are separated so that they can be reused together with other tables, such as caregiver information.
6.4.2 Caregiver information

The advantages of separating the address fields to own tables can be noticed in Figure 11. Table `regAddress` can be used to save the address information related to both Client and Caregiver. Basically, it can be used wherever address information has to be saved irrespective of the entity.

---

*Figure 10 Database relationship diagram for storing user’s basic information at server.*
6.4.3 Device information

A client may use many devices, so the relation between Client and Device table is one-to-many (1: M). Every device is identified by a MacId. Device information can be stored in regDevices table.

Figure 11 Database relationship diagram for storing caregiver information at server.

Figure 12 Database relationship diagram for storing User’s device information at server.
6.4.4 Risk factors

Whenever a user enters the information into the app and requests the risk score, the entered information related to risk factors will be saved in `regRiskFactors` table. The relation between Client and Risk Factors table is 1: M, meaning a client can have many records of risk factors with a different time stamp. In this way, it will be easier to follow up on the Client risk scores for a certain time period. As the information will be saved server-side, it can be used for recommendations to other patients who has the same symptoms, or who falls under the same category.

6.4.5 Risk score information

Figure 13 Database relationship diagram for storing user’s risk factors with different timestamps at server.

Figure 14 Database relationship diagram showing relations between user, caregiver and users calculated risk score with different timestamps.
7 Conclusion

7.1 Discussion

The latest CAIDE study results include many new factors which were not in previous studies. Based on preliminary results, there are risk factors such as DBP, the number of heart diseases, and loss of memory that may play a significant role in total score calculations. If the app is used by caregivers, then there is higher probability that they can check the previous medical records of the person and take the mean value of DBP in recent days, for example, instead of current DBP (which varies quite a lot from time to time) and they can even check if the person is suffering from any heart conditions. In order for the app to check such records automatically the medical database owners should provide some standardised open API that the app can be authorised to use if motivated. The CAIDE app maker can define the requirement specification for the API (i.e. what the input parameters they can send are and what the output should look like) and make a standardised configuration (e.g. server name and credentials to access API) which each caregiver should fill out in order to integrate the app with their own databases. Who will provide the money for caregivers to implement such an API is out of context here but for the Swedish medical system, it should be enough to develop the API as a standard template and make it configurable for each caregiver, as the key will be Swedish personal number all over Sweden.

Figure 15 Sequence diagram: Retrieval of DBP from Caregivers Database for Risk score calculation.
If the above model is used by the caregiver, then it does not even need a connection to the server. It needs only connection to the API Server which will be configured individually by the caregivers themselves. There is no need to expose the app to the Internet, to avoid risks related to security. Above shown is a sequence diagram which explains what a request and response will look like while retrieving the mean DBP recorded at the caregiver in the last 48 hours (just an example).

There can be a few factors that can influence the calculated risk score but have nothing to do with the person for which the risk score is being calculated. For example, if the application user is not the person for whom the risk score is being calculated, like when a daughter is calculating the risk score for her elderly mother. Here are the examples of possible errors in input data. (A) If the app user is not the person for whom the risk score is calculated, then there is a higher probability for erroneous input data. The user might feed false information to app purposely or by mistake, in any case, it would be very difficult to detect such false information. (B) There can be bugs in the app where the user feeds some information into the GUI but while saving the information the parser might have misrepresented the information to be recorded in the database. For example, a decimal is saved as an integer or a string is truncated while saving to database. (C) Mixing different types of units can give false risk score, for example, blood pressure in mm Hg instead of kPa (kilopascals), or height in feet instead of meters. (D) If the person has a tendency to answer always ‘yes’ for questionnaire then the result of calculated risk score effects. The risk score app should be able to determine if a person has a syndrome of saying always “yes” or “no”, for example, by analysing data from other apps and comparing it with the input data. If the person has such tendency then the questions used in the app can either reframed slightly different or can be divided into several questions instead of one. An example is, instead of asking a question like “Are you physically active?” to determine if a person is physically active or not, Several questions can be asked like “How long do you think you walk per day?”, “How long do you run per day?”, “How many days in a week do you swim?” or “How many hours per week do you spend your time in Gym?”. So instead of answering a direct question whether the person is physically active or not is determined by a set of questions. The wrong answer need not be always an indication of yes or no syndrome, but it can be the result of the complexity of queries where either user misunderstood the question or not understood the question at all. Another way of correcting the input data could be to check data from another app on user’s device for example in this case from a pedometer and compare the data with user’s input.

The calculated risk score may be affected by application-related settings, e.g. if the app collects data from other installed programs in order to calculate the calories burn per day; the app can calculate how many steps a person took in a day or how many meters the person climbed, etc. If the person forgot the mobile at home while walking or if he or she lent his mobile to his friend who is a marathon runner, then false information is collected by the app without the user knowing about it.

Mobile app usage increases day by day, as discussed in Chapter 2.3. If we consider mental-health apps, there are a number of apps developed in recent years [80]. Due to unavailability of data, among other factors, many people are not getting the correct treatment for mental disorder. Mobile health apps can help to fill the gap. Mobile apps are becoming ubiquitous because the user can use the app at anytime and anywhere as these devices are light and easy to carry. Apps may or may not be effective, and some may even be harmful, e.g. Promillekoll, a smartphone app created by Sweden’s government-owned liquor retailer, designed to help curb...
risky drinking. Using Promillekoll, the user enters each drink they consume during a party and the app returns an approximate blood-alcohol concentration. When Swedish researchers tested the app on college students, the randomly selected men ended up drinking more frequently than before, although total alcohol consumption did not increase. Here app users may have felt that they can rely on the app to reduce negative effects of drinking and therefore ended up drinking more. This may be due to the user of the app not knowing how to use the app, so the app users should learn how to use the app and test it few days before starting to use it for the purpose it is meant for. A lot of data related to medical and non-medical factors is already available around the world among different research groups which could be very helpful in predicting dementia and Alzheimer’s disease. Integration of all this data is a challenging task which needs extensive co-operation among the research groups and countries. Many decisions might even need political approvals as well which could be very time-consuming and it could be very difficult to get finance for such wide research projects.

As described in Chapter 3, genomic data particularly related to APOE*E4 is important in determining AD risk. But there is limited data related to genomes is available for research due to privacy issues. There is a need to conduct programs, policies to increase the awareness related to sharing genomic data and about precision medicine in general people so that research results into products, the knowledge that improves the efficiency of predictability tools. In context to this, National Institutes of Health (NIH) announces the final Genomic Data Sharing (GDS) Policy that promotes sharing of genomic related data for research purposes. Future apps can have the option to enter genome related data to predict the future disease risk and can control the lifestyle-related factors as genomic data can be available for the common man (due to a decrease in genome costs (Figure 3) and increase in awareness of genomic related data).

**Limitations**

There are a lot of non-medical risk factors that may have an effect on dementia. But this study resulted in identifying only some of them. Instead of introducing all identified risk factors into risk score calculation, one risk factor from each category is discussed and due to unavailability of data, it was hard to specify the risk factor weight in final risk score. From discussion related to personal genomics only APOE is considered in risk score but the effect of other genes such as CNU, CR1, PICALM, BIN1 etc. were not considered as these may not have the importance as APOE*E4 gene in risk of Alzheimer’s disease.

**General aspects**

The history related to some of the discussed risk factors may plays an important role in calculating final risk score. For example considering SBP which varies frequently, a history data related to SBP is more helpful while calculating final risk score instead of considering current SBP. For example a person may have high SBP during last 10 years and may have lower SBP currently that does not mean person’s risk related to dementia decreased. The earlier SBP may plays a role in dementia at late-life. The same applies for smoking and air pollution. Along with current risk factor data, history of risk factor data plays a role in final risk score.

Amount of intake related to risk factor is also plays a role in final risk score. For example moderate alcohol consumption may not affect the final dementia risk but high intake of alcohol may increase the dementia risk. Similarly moderate coffee drinking (8 cups per day) may not increase dementia risk instead it helps in decreasing the dementia risk.
The risk score enhancement achieved using ICT-tools should be validated by clinicians and caregivers, for example before installing the app on the client’s device, the terms and conditions should clearly mention how the app will collect information from other apps, and type of information that will be saved at server-side. It is a good idea to provide an option for the user as to whether the app is permitted to save data at the server side or not. The user should also be able to easily find out how the saved data will be used and who will be having access to such data. For instance, once per year the user should have a possibility to get a copy of the data collected. The app should receive the user consent before publishing the results to any third party, and there should be options for the users to share the data (anonymous or non-anonymous).

All the sensitive information that flows between app and server should be encrypted and the information should be saved to a secured server. If the information is to be used by any other researchers, there should be a written agreement between the researching party and the database owner so that the information will not be further distributed unnecessarily.

The risk score from the app should be mostly used as information by users and caregivers. Caregivers should not prescribe any medication solely based on risk scores calculated by the app. Clinical diagnosis should always be performed before prescribing any medication. The app should never let the client come under the impression that he or she is suffering from dementia or AD.

Although most of the research is done by non-profitable organisations such as SICS, KI/ARC, etc. which in this case are financed by the European Union to serve its citizens, the app still needs a sustainable source of income for the maintenance and further development. There will also be running costs for the server and database. There are also costs related to combining medical and non-medical data. One idea could be introducing advertisements for the generation of revenue, but the risk is that the user can get an impression that the app is commercial and that their personal integrity, therefore, can be under threat. Other possibilities could be that all the inventions made under the research can be patented by the research parties and the income generated by these patent rights could be used for the app’s survival and refinements.

### 7.2 Concluding remarks

As discussed in Chapter 1.3 the goal of the thesis is to contribute to an increase in the usefulness of ICT tools that predict cognitive decline in elderly people. This is possible by increasing the underlined risk score used by these prediction tools. The following efforts were made to increase the risk score efficiency.

1) Understand the role of ICT in producing predictive tools for dementia, with a focus towards individual treatments (Chapter 2 and Chapter 3)
   Chapter 2 includes a discussion about the role of ICT in active ageing, explains various risk factor categories. Chapter 3 describes the new technologies that were targeted towards the individual i.e. genetically related factors. This section describes the terms precision medicine and personal genomics.
   The chapter helps in understanding the current ICT trends in developing tools and describes the focus of treatment from general medicine to precision medicine, which is important in considering the new development of risk score tools for ex. dementia prevention tools.

2) Understand existing risk scores (Chapter 4)
Various risk scores are described in Chapter 4, particularly the old CAIDE risk score, the basis for the new CAIDE risk score, which in turn is planned for use in a new evidence-based tool (app). Other late-life risk scores have been discussed to understand the risk factors involved and the importance of these risk factors.

3) Identifying non-medical risk factors (Chapter 5)
   The importance of non-medical factors in calculating risk scores is discussed in Chapter 5. An attempt is made to find the additional non-medical risk factors which can be included in future risk score calculations for effective risk score results.

4) Suggested design for filling the gap between data available at server and introducing this data in risk score calculation (Chapter 6)
   The old CAIDE risk score algorithm, which was not previously described in the published literature, was presented. Effects of considering non-medical factors in the risk score calculation are discussed with example risk factor from each discussed category and showed that risk score has been enhanced when compared to previous. A simple database design (tables, fields, relationships) that can be used to store the information at server side is also presented, and an app architecture at server side is discussed. How one can improve the risk score by adding and validating some non-medical data discussed in Chapter 5, and showed that there is a change in risk score by considering them. The suggested architecture can be a prototype for a tool currently being developed by researchers at SICS, as part of MULTI-MODE.
References


24. My-AHA, My Active and Healthy Ageing, Technology in use. [http://www.activeageing.unito.it/node/19](http://www.activeageing.unito.it/node/19) [Last visited: 2016-08-07].


40. Miia Kivipelto, Tiia Ngandu, Tiina Laatikainen, Bengt Winblad, Hilkka Soininen, Jaakko Tuomilehto. Risk score for the prediction of dementia risk in 20 years among middle


Appendix A

Gerontechnology provides an enhanced environment for adults by combining technology and ageing. It involves the research and development of techniques, technological products, services, and environments based on knowledge of ageing processes [1]. Gerontechnology helps older people to lead healthier lives, more independently, and socially engaging ones. Gerontechnology goals are classified into prevention, compensation, care, and enhancement [2].

Prevention: Long-term intervention trials are proposed to prevent or delay the age-related declines. For example activities like exercise; diet helps to slow down physical, mental and social decline while ageing. Improving diet by using a nutrition tool, stay fit and keep walking by following training plan using exercise videos, playing computer games to simulate brain, social robots whose tasks is to engage older adults in natural social exchanges or to provide some form of interaction so they do not feel loneliness are the examples of ICT prevention methods.

Compensation: Compensation comes into picture when disability is no longer prevented. To reduce the impact of disability or to replace physical activities for elder people ICT is playing an important role. Actuators or sensors are helpful in physical activities like closing doors to turning on/off lights, to open/close fridge doors. Sensors can be used to know the bed/chair occupancy, electrical usages. The data provided by these sensors can be available to caregivers/family members to help older ones if necessary.

Care: Care comes into picture when a problem cannot be compensated or prevented. Emergency alarms, sensors can be used for the better care of elderly people. Medical emergency response systems/emergency alarms are used for fall detection. In this system, senior wears a transmitter around the neck, on their wheelchair, wrist. When help is needed, the senior presses the transmitter button to reach caregivers. Installation of sensors like smoke, fire detectors at elderly home are helpful in providing the necessary help.

Enhancement: Technology helps in creating new opportunities and extending existing capabilities. Using IT to keep updated with our interests, to find interesting activities to join, to learn new things, to get new social opportunities and friendship using social networks, to enable real-time access to information, to enable work from home opportunities are the examples of IT enhancement.
References for appendix
