Annamari Lundqvist
Tomi Mäki-Opas
(eds.)

Health 2011 Survey – Methods

REPORT

8 | 2016

REPORT 8/2016

Annamari Lundqvist and Tomi Mäki-Opas (editors)

Health 2011 Survey – Methods

TERVEYDEN JA
HYVINVOINNIN LAITOS
Institutet för hälsa och välfärd

© Authors and National Institute for Health and Welfare

Cover drawing: Riitta Nieminen
Layout: Annamari Lundqvist

ISBN 978-952-302-668-1 (printed)
ISBN 978-952-302-669-8 (online publication)
ISSN 1798-0070 (printed)
ISSN 1798-0089 (online publication)
http://urn.fi/URN:ISBN:978-952-302-669-8

Juvenes Print – Finnish University Print Ltd
Tampere, Finland 2016

Foreword

The target group of this report is researchers planning to use the Health 2011 Survey data for various health monitoring and scientific research purposes, and those who are planning and conducting large population based health examination surveys in the future. The aim is to enhance research collaboration at the international level and therefore the report has been published in English only.

We wish to thank all key experts contributing to the planning and implementation of the Health 2011 Survey. Their valued expertise and efforts laid the basis for this methodology report. However, only persons participating in the drafting and writing this methodology report are listed as authors.

We also would like to thank:

- all the 7,488 subjects who gave their time to take part in the Health 2011 Survey,
- the fieldwork personnel working so diligently to collect the data,
- Research Professor Seppo Koskinen who lead the Health 2011 Survey and contributed to planning and reviewing this report,
- Research Manager Päivikki Koponen who was involved in the planning and implementation of the survey and planning, drafting and critically reviewing this report,
- Professor Arpo Aromaa who had a very important role in planning the survey and reviewing this report,
- Dr. Sebastián Peña who made an invaluable contribution to the planning the survey,
- RN Ulla Laitinen and MSc Minna Collins-Cona for their careful and devoted editorial assistance,
- the National Institute for Health and Welfare for providing excellent research facilities and support, and
- the numerous agencies who contributed to the funding of the survey.

We are also grateful to all others not specifically mentioned but who contributed to the planning and implementation of the survey as well as the preparation and publication of this report.

Abstract

Annamari Lundqvist and Tomi Mäki-Opas, eds. Health 2011 Survey – Methods. The National Institute for Health and Welfare (THL), Report 8/2016, 219 pages, Helsinki 2016, ISBN 978-952-302-668-1 (printed), ISBN 978-952-302-669-8 (online publication).

This report describes the planning, design and implementation as well as the main methods and contents of the Health 2011 Survey, a comprehensive national health examination survey on the levels, changes and determinants of health, functional capacity and welfare of the population. Guidelines are also included for using Health 2011 data for research purposes.

The main aim of the Health 2011 Survey was to provide an up-to-date overview of health and functional capacity in the working-age and elderly populations as well as time trends in them. In the survey, the prevalence of the most common health problems and factors contributing to them as well as need for treatment, rehabilitation and assistance associated with these problems were evaluated. The survey design is both cross-sectional and longitudinal.

The invitation to take part in the Health 2011 Survey was sent to all persons who had been included in the representative, two-stage stratified, random sample of the national Health 2000 Survey 11 years earlier (N=8,135). A new random sample of persons aged 18–28 years was also drawn (N=1,994). In addition, 920 subjects participating in the Mini-Finland Survey in 1978–1980 and who were invited to a resurvey in 2001 were invited.

The Health 2011 Survey was carried out by the National Institute for Health and Welfare (THL) in co-operation with researchers and experts from several other organisations. The fieldwork was carried out in 2011–2012 by five teams in 60 locations all over the country. A wide range of information was collected by a comprehensive health examination, face-to-face and phone interviews as well as self-administered questionnaires. In addition, register based information on all people in the sample was linked to the survey data.

Nearly three quarters, 73.6 per cent of the subjects aged 29 years or older took part in at least one data collection phase. The participation rate was 42.3 per cent in the new cohort of young adults and 81.3 per cent among the Mini-Finland resurvey sample. First results of the survey were published in 2012.

The Health 2011 Survey data provide exceptionally good opportunities for health monitoring as well as for multidisciplinary public health and epidemiologic research. Register based follow-up also enhances the possibilities for scientific research.

Tiivistelmä

Annamari Lundqvist ja Tomi Mäki-Opas, toim. Terveys 2011 -tutkimus – Menetelmät. Terveyden ja hyvinvoinnin laitos (THL). Raportti 8/2016, 219 sivua. Helsinki 2016, ISBN 978-952-302-668-1 (painettu), 978-952-302-669-8 ISBN (verkkojulkaisu).

Tässä raportissa kuvataan Terveys 2011 -tutkimuksen suunnittelu ja toteuttaminen sekä tutkimuksen sisältö ja menetelmät. Raportti käsittelee myös tutkimusaineiston hyödyntämistä ja suositeltavia tilastollisia analyysejä.

Terveys 2011 -tutkimuksen päätavoitteena oli tuottaa kattava ja ajankohtainen kuva väestön terveydestä, toimintakyvystä ja niiden muutoksista. Tutkimuksessa selvitettiin tärkeimpien terveysongelmien yleisyyttä ja syitä sekä niihin liittyvän hoidon, kuntoutuksen ja avun tarvetta. Tutkimuskysymyksiä on mahdollista tarkastella sekä poikkileikkaus- että pitkittäisasetelmassa.

Terveys 2011 -tutkimukseen kutsuttiin kaikki elossa olevat Terveys 2000 -tutkimukseen 11 vuotta aiemmin kutsutut henkilöt (N=8 135), jotka edustavat Suomen aikuisväestöä. Lisäksi poimittiin uusi täydentävä satunnaisotos (N=1 994) 18–28-vuotiaita nuoria aikuisia. Tutkimukseen kutsuttiin myös ne 920 henkilöä, jotka olivat osallistuneet Mini-Suomi -tutkimukseen vuosina 1978–1980 ja jotka oli kutsuttu uusintatutkimukseen vuonna 2001.

Terveys 2011 -tutkimuksen suunnittelua ja toteutusta koordinoi Terveyden ja hyvinvoinnin laitos (THL) yhteistyössä laajan asiantuntijaverkoston kanssa. Tutkimuksen kenttävaihe toteutettiin vuosina 2011–2012. Tiedonkeruusta vastasi viisi kenttäryhmää, jotka työskentelivät 60 tutkimusalueella eri puolella Suomea. Tutkimuksessa kerättiin monipuolisesti ja laaja-alaisesti tietoa terveystarkastuksen, haastatteluiden ja kyselylomakkeiden avulla. Tutkimusaineistoa on täydennetty yhdistämällä siihen kansallisista rekistereistä saatavia tietoja.

Tutkimukseen kutsutuista 29-vuotiaista ja sitä vanhemmista henkilöistä 73,6 prosenttia osallistui ainakin yhteen tiedonkeruun vaiheeseen. Nuorista aikuisista tutkimukseen osallistui 42,3 prosenttia ja Mini-Suomi-otokseen kuuluvista 81,3 prosenttia. Ensimmäiset tutkimustulokset julkaistiin vuonna 2012.

Terveys 2011 -tutkimuksen aineisto tarjoaa ainutlaatuisen mahdollisuuden sekä väestön terveyden seurantaan että monitieteelliseen kansanterveyden tutkimukseen. Rekisteriseurantatiedot otokseen kuuluvien terveydestä lisäävät oleellisesti aineiston käyttökelpoisuutta korkeatasoiseen tieteelliseen tutkimukseen.

Sammandrag

Annamari Lundqvist och Tomi Mäki-Opas, red. Hälsa 2011 – Metoder. Institutet för hälsa och välfärd (THL). Rapport 8/2016, 219 sidor. Helsingfors 2016, ISBN 978-952-302-668-1 (tryckt), ISBN 978-952-302-669-8 (nätpublikation)

Den här rapporten beskriver planering, design, utförande och metoder i Hälsa 2011, en nationell undersökning om befolkningens hälsa, funktionsförmåga och välfärd samt om bidragande faktorer och förändringar i dem. Rapporten innehåller också rekommendationer rörande analysmetoder.

Det primära syftet med undersökningen Hälsa 2011 var att ta fram en heltäckande aktuell bild av hälsan och funktionsförmågan samt förändringarna i dem. Undersökningen har kartlagt förekomsten av de viktigaste hälsoproblemen, orsakerna till dem och det anknytande behovet av vård, rehabilitering och hjälp. Undersökningen är till sin design både en tvärsnitts och en longitudinell studie.

Till undersökningen Hälsa 2011 kallades de ännu levande personer som 11 år tidigare hade ingått i det representativa stickprovet Hälsa 2000 (N=8 135). Därtill undersöktes ett antal (N=1 993) slumpmässigt utvalda 18–28-åringar och 920 människor som hade deltagit i Mini-Finland undersökningen åren 1978–1980 och som hade inbjudits till undersökningen igen år 2001.

Hälsa 2011 var ett omfattande samarbetsprojekt som koordinerades av Institutet för hälsa och välfärd (THL). Fältarbetet utfördes under åren 2011–2012 av fem undersökningsteam som arbetade inom 60 områden täckande hela Finland. Omfattande information samlades vid en mångsidig hälsoundersökning, inkluderande intervjuer och frågeformulär. Därtill skaffades information från olika nationella register.

Av de personer som var 29 år gamla eller äldre deltog 73,6 procent i något skede av undersökningen. Deltagande var 42,3 procent bland yngre personer (ett nytt urval) och 81,3 procent bland personer som tillhört stickprovet tagit av Mini-Finland. De första resultaten publicerades år 2012.

Det insamlade informationen erbjuder sällsynt goda mjöligheter för hälsouppföljande och tvärvetenskaplig forsking inom epidemiologi och folkhälsovetenskap. Uppföljningen via register främjar också högklassig vetenskaplig forsking.

Contents

Introduction

Annamari Lundqvist, Päivikki Koponen, Tomi Mäki-Opas and Seppo Koskinen

In order to promote health and welfare, to plan services and to develop policies, a clear picture of the status and development of the population's health and welfare is required. In Finland, information on population health has been collected by health examination surveys since 1965. From 1965 through to the 1970s, the Social Insurance Institution (KELA) implemented mobile clinic health examination surveys in different parts of the country (Heinonen 1966). The National Institute for Health and Welfare (THL, the former National Public Health Institute, KTL), for its part, has conducted FINRISK surveys in several parts of the country at five-year intervals (Borodulin et al. 2015b), focusing mainly on cardiovascular diseases and their risk factors. The first comprehensive combination of a health interview and health examination was the national Mini-Finland Survey that was carried out between 1978 and 1980 (Aromaa et al. 1989). The Health 2000 Survey with a representative sample of Finnish adults and an extensive data collection on many major public health problems was conducted between 2000 and 2001 (Aromaa and Koskinen 2004, Heistaro 2008). A follow-up survey, the Health 2011 Survey, provides up-to-date information on the most important public health problems in the country, their causes and treatment as well as on the population's functional and working capacity. As many methods were chosen with a specific view to ensuring comparability with earlier studies, comparisons of the new Health 2011 data with those from the Mini-Finland and the Health 2000 Surveys were assured to identify changes in health, welfare and functional capacity in the population and to provide a solid foundation for prediction of future trends. The study builds a basis for health monitoring, health policy, planning and evaluation of health promotion and services, multidisciplinary public health and epidemiologic research.

The Health 2011 Survey was led by THL. The project organisation involved a large number of partners and experts from different universities, research institutes and other organisations. The planning of the Health 2011 Survey was begun in 2009 and the fieldwork was conducted in 2011–2012. Key results were published in 2012 (Koskinen et al. 2012).

This report describes the planning, implementation and main contents of the Health 2011 Survey as well as the guidelines on the use of the data for research purposes. In addition to the material published in this report, additional material such as various questionnaires used in the survey, is available on the Health 2011 website (www.thl.fi/terveys2011).

HEALTH 2011 SURVEY – PLANNING, IMPLEMENTATION AND DATA

1. PLANNING AND PREPARATION

1.1 Project organisation and funding

Seppo Koskinen, Annamari Lundqvist and Tomi Mäki-Opas

Project organisation

The project organisation involved a wide range of organisations (see Appendix 1). THL had the overall responsibility for the project planning and implementation. A large number of specialists from different organisations participated in the project organisation.

The advisory board was chaired by the Director General of THL, Professor Pekka Puska, and the members represented the major collaborators and funding organisations. A steering group under the chairmanship of Professor Marja Vaarama, THL, was in charge of the general direction of the project. The management group, chaired by Professor Seppo Koskinen, THL, was responsible for the planning and execution of the Health 2011 Survey. Much of the detailed planning of the fieldwork was carried out in the co-ordinating group of the fieldwork, led by Dr. Tomi Mäki-Opas. Expert groups on different main topics of the survey participated in all phases of the preparation and execution of the survey. Among the topics covered by these teams were living environment, cardiovascular diseases and diabetes, mental health, musculoskeletal disorders, oral health, reproductive health, health behaviour, functioning as well as need and use of services.

The project organisation involved more than 200 researchers and other experts covering different topics and later taking part in planning, training and supervision of the field examinations, as well as reporting. Researchers both from the research institutes involved in the project and from a number of universities and hospitals have been involved in reporting studies based on the data.

Funding

The overall costs of the preparation of the fieldwork as well as collecting and editing the data and preparing the baseline report totalled approximately 3.5 million euros, not including the large amount of work carried out by permanent experts at THL and other participating organisations. The contribution of the permanent personnel to the project was part of their normal work. The funding was collected from nearly twenty sources. The largest contribution was received from the KELA, and the other main sponsors were the Ministry of Social Affairs and Health, the UKK Institute, the Local Government Pensions Institution, the Finnish Dental Association and the Finnish Dental Society, the Academy of Finland, the Finnish Centre for Pensions, the Federation of Finnish Financial Services, the Finnish Institute of Occupational Health and the Finnish Work Environment Fund. In addition, smaller contributions were received from different universities and other organisations.

1.2 Sampling design

Tommi Härkänen and Esa Virtala

In the Health 2011 Survey, the study subjects of the Health 2000 Survey were re-invited in order to form a representative longitudinal data on the Finnish population. In addition, a new sample of young adults was taken. The sample sizes are presented in Table 1.2.1 and Figures 1.2.1 and 1.2.2.

Sampling design of the Health 2000 Survey

The target population of the Health 2000 Survey comprised individuals aged 18 years or older and living in mainland Finland on 1 July 2000 (Laiho 2004). In addition to the household population, people living in institutions were included. The Autonomous Territory of Åland Islands was excluded, as were people living on islands not accessible by road.

A stratified two-stage cluster sampling design was used. Mainland Finland was divided into 20 strata defined by the 15 largest cities and towns (their health centres) and the remaining rural areas based on the five university hospital regions. The 15 towns were selected with probability

1, and the remaining 65 health centres were selected from the rural strata using a systematic probabilities proportional to size (PPS-SYS) design. The second stage involved sampling individual persons from those districts. The sample sizes for each health centre within a stratum was equal so that the total sample size in a stratum was proportional to the target population. Oversampling of the people aged 80 and older was carried out using double inclusion probabilities. The total sample size was 9,922. Of these, those who were at least 30 years of age (N=8,028) were invited to participate in the health examination; young adults (N=1,894) were invited to participate in the health interview and to fill in the questionnaires (Härkänen et al. 2016).

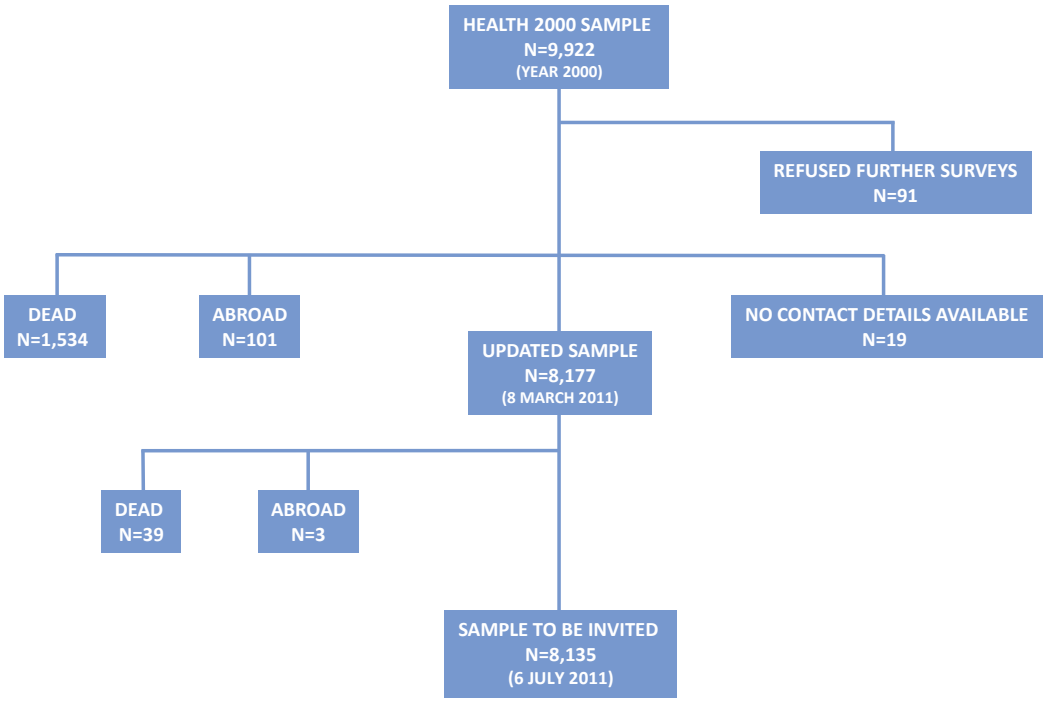
Sampling design of the Health 2011 Survey

All members of the Health 2000 Survey sample, who were alive, living in Finland in 6 July 2011, had contact details available and had not refused to participate in further surveys were invited to take part in the Health 2011 Survey (N=8,135, Figure 1.2.1). Study subjects were at least 29 years of age.

Table 1.2.1. Sample sizes in the Health 2011 Survey.

Sample	Age group (yrs)	Men	Women	All
Health 2000				
	41–50	820	875	1,695
	51–60	826	899	1,725
	61–70	712	750	1,462
	71–80	388	513	901
	81–	152	384	536
	All	2,898	3,421	6,319
Health 2000 young adults				
	29–34	484	451	935
	35–40	460	421	881
	All	944	872	1,816
New sample of young adults				
	18–23	580	537	1,117
	24–28	436	441	877
	All	1,016	978	1,994

Figure 1.2.1. Formation of the Health 2011 sample.



Representativeness of the sample

The baseline sample represents the target population of each stratum on 1 July 2000. Mortality and emigration of the individual sampling units represent the mortality and emigration of the population, because no intervention was conducted on the sampling units. Migration of the study subjects between the strata was also representative. The target population now was people, who belonged to the baseline target population and were alive and living in Finland in 2011. However, immigration to Finland after the year 2000 has not been represented by the fixed panel defined by the baseline sample. However, a separate health examination survey on three major migrant groups was carried out at the same time and with comparable methods as the Health 2011 survey (Castaneda et al. 2012).

New sample of young adults

To cover also young adults, a new random sample of persons aged 18–28 years was included in the Health 2011 Survey (N=1,994). Of these, 415 subjects who belonged to the COURAGE and Physical Activity and Fitness sub-studies (see below and Chapters 19.1 and 19.2) were asked to participate in the health examination. A self-administered questionnaire was sent to 1,579 subjects.

Due to the fact that there were changes in the health centre districts from year 2000 to 2011, a modification for the original PPS based sampling design was required. In three cases where old health centre districts were split and some of the municipalities were amalgamated to another municipality (Piikkiö to Kaarina) or to a larger town (Korpilahti to Jyväskylä, and Savonranta to Savonlinna), a large sample was sampled from the new town and the GIS coordinates were utilized to select an appropriate number of individual sampling units from the amalgamated municipalities. Two health centre districts were amalgamated to other health centre districts (Pyhäselkä to Joensuu, and Perniö to Salo).

Furthermore, there can be overlap or underlap between the true municipal borders and the polygon-based, approximate municipal borders, but we consider these differences small, because the underlap or overlap areas were usually sparsely populated. The population sizes in 2011 based on the municipal boundaries in 2000 were obtained from Statistics Finland.

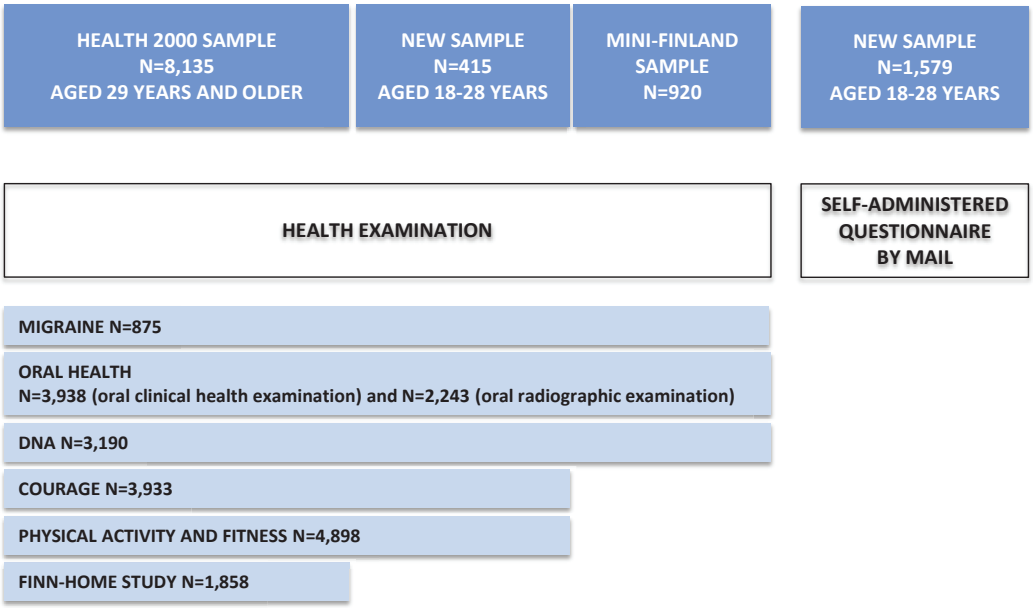
Sample of the Mini-Finland resurvey

A total of 1,278 participants of the Mini-Finland Survey (Aromaa et al. 1989) were invited to participate in the Mini-Finland resurvey conducted in seven municipalities in 2001 (Heistaro 2008). Those 920 subjects, who were alive in 2011, were invited to participate in 2011–2012 (Figure 1.2.2). In 2001, there were some differences in the survey protocol between the samples of Mini-Finland resurvey and Health 2000 Survey (Heistaro and Koskinen 2008) while a similar survey protocol was applied for both samples in 2011.

Samples for sub-studies

The samples for sub-studies are illustrated in Figure 1.2.2.

Figure 1.2.2. Sample sizes in main Health 2011 Survey and in sub-studies.



New sub-studies, the COURAGE and the Physical Activity and Fitness, were based on random subsets of the full Health 2011 Survey sample. The sample sizes of these studies were 3,933 and 4,898, respectively. It is worth to note that among 29–74 years old, the COURAGE and the Physical Activity and Fitness samples did not overlap, but in the age groups 18 to 28 years, and 80 years and above, the overlap was complete.

The clinical oral health examinations were carried out in the southern and northern field examination areas (i.e. Helsinki and Oulu) and the oral radiographic examination was performed only in Helsinki. A total of 3,938 subjects were invited to participate in clinical oral health examinations and 2,242 subjects were invited to participate in the oral radiographic examination.

In the Health 2000 Survey, all subjects who were 30 years of age or older were asked to provide a blood sample for DNA extraction. In the Health 2011 Survey, those subjects who did not participate in the Health 2000 Survey or did not have a blood sample available for genetic analyses were part of the DNA sub-study (N=3,190).

A total of 2,120 persons aged 45–74 years old were invited to participate in the home blood pressure measurement sub-study called Finn-Home study in 2000–2001. All those subjects were invited to take part in home blood pressure measurement also in Health 2011 Survey (N=1,858).

Subjects who reported to suffer from severe (moderate or intense) headaches in the health interview were invited to participate in a sub-study regarding migraine (N=875).

1.3 Ethical approval

Tomi Mäki-Opas

The plans and protocols of the Health 2011 Survey were submitted for approval to the relevant ethical committees. The application was first reviewed by the Ethical Committee of THL. A more detailed project plan was submitted to the Coordinating Ethics Committee at the Hospital District of Helsinki and Uusimaa (HUS, reference 45/13/03/00/11). At both these stages, the plans received favourable opinions. However, the ethical committee round in HUS Coordinating Ethics Committee took four rounds while improvements and corrections were asked to be submitted to the research plan, covering letter, information letter, information to the research subject and to the copy of the statement.

The study protocol on radiographic oral health examination was approved by the Advisory Board for Radiation Safety and the safety license was granted by the Radiation and Nuclear Safety Authority, Finland.

1.4 Preparation of fieldwork

Tomi Mäki-Opas, Päivikki Koponen, Annamari Lundqvist, Ulla Laitinen, Pirkko Alha and Seppo Koskinen

The planning of the Health 2011 Survey started in 2009. The fieldwork protocol including questionnaires, interviews, and the health examination protocol was designed by the Health 2011 project team in THL together with several key experts from THL, Finnish Institute of Occupational Health, University of Helsinki and several other agencies (see Appendix 1). Validated and widely used methods were chosen whenever possible to maximize the quality of data and international comparability of the results. Moreover, the methods and contents of survey were aimed to be as similar as possible with the Health 2000 Survey to ensure comparability of the results across the surveys. However, due to the financial restrictions, not all measurements were included in this most recent survey (e.g. clinical medical examination was not included in the Health 2011 Survey). The methods and contents of the survey are described in the following chapters. All the questionnaires used in the survey are available in the website of Health 2011 Survey (www.thl.fi/terveys2011).

Recruitment of fieldwork personnel

A total of 85 fieldwork nurses (including a few bioanalysts/laboratory technicians and physiotherapists, all later called as nurses) from the five main regions in Finland (Helsinki, Kuopio, Tampere, Turku and Oulu) were recruited. In addition, a dentist and a dental nurse were needed for each fieldwork team. The fieldwork personnel were first recruited through the internet service of the national recruitment office. As not enough applicants were found, job advertisements were placed in the main regional newspapers. After several rounds of interviews, all positions were filled with the exception of three dentists and three dental nurses. Due to the lack of personnel, only the fieldwork teams in Oulu and Helsinki comprised dental personnel and performed clinical oral examinations as a part of the health examination.

Recruitment of competent nurses turned out to be a challenge due to the general shortage of nurses affecting the entire health care system in Finland. Among the fieldwork personnel recruited, there were some nurses with previous experience of fieldwork in health surveys while most of the nurses were not familiar with health surveys. During the training period and the first weeks of fieldwork, some nurses were dissatisfied their salaries and travel allowances

and some were found not to be competent for health survey fieldwork tasks. This led to some turnover of personnel, and five new nurses were recruited later during the fieldwork.

Pilot survey

A three-day pilot survey was conducted in May 2011, at the main office of THL in Helsinki. This was not a full pilot, but rather a pre-testing of the survey protocols and timing with a small number of participants.

Altogether 38 volunteers of different sex, age, and social background were recruited from the family and friends of the research team. The participants in the pilot study received the same materials and questionnaires that individuals in the actual study sample were going to receive.

The key experts and the research team of the survey participated in the pilot study by working as fieldwork personnel and by observing the measurements as well as the feasibility and fluency of the survey protocol as a whole, and collecting feedback from the participants.

Voluntary participants of the pilot study were asked to fill in a feedback questionnaire inquiring about their experiences of survey protocol. Feedback was also given by key experts and research personnel involved in the pilot study.

The following issues were considered after the pilot:

- The survey methods were mainly developed on the basis of experiences from the previous survey in 2000 and other surveys carried out at THL. As many issues in THL organisation, in public attitudes and the population had changed, a more detailed pilot would have been needed to further specify the recruitment, measurement and training procedures. This would have facilitated planning of the training of the fieldwork personnel and finalising the survey protocols.
- It was not possible to evaluate the population's willingness to participate in the study and the recruitment process of the fieldwork personnel because the participants were volunteers and fieldworkers were recruited from the research personnel.
- The evaluation of the survey protocol was limited by the fact that not all survey materials were ready before the pilot. Recorded timing and average duration of interviews and examinations per participant were used to evaluate how realistic the estimated and budgeted personnel resources were, and the potential burden to participants. Based on the experiences of voluntary participants and personnel, some

adjustments for the fieldwork protocol were made. Especially, the questionnaires and protocols were adjusted to better meet the restricted resources.

- The use of equipment, computer programs and data management should have been tested in more detail. The processing, transfer and storage of blood samples are routine for THL where several surveys are in the field every year, so there were no specific needs to evaluate this in the pilot. The fact that not all equipment and computer programs were finalised before the pilot caused some problems and several updates for the computer programs were made during the first two weeks of the survey. This caused not only problems in transfer and quality of data but also some frustration among the fieldwork personnel.
- The aim of the pilot was to identify potential practical problems. Some practices were refined and further specifications were added to the manual, but some changes and additional instructions needed to be made even after the fieldwork had been started, as the small pilot had not revealed all problems which became evident after the fieldwork begun.
- Although the pilot survey was carried out three months before the actual survey, there was not sufficient time to evaluate the pilot in detail and to make many adjustments to procedures before the full-size survey. This was mainly due to the summer holiday season between the pilot and the full-size survey.
- The pilot survey should have been conducted by the same personnel who were recruited for the full-size survey, so that each team would have had at least one person with experiences from the pilot. This was not possible due to the restricted budget.

Piloting of the oral clinical examination was separately carried out at a public dental clinic in April 2011 (see Chapter 13). The pilot of the COURAGE study is described in Chapter 19.2.

Training

Two-week training for all fieldwork personnel was organized between 25 June and 5 August 2011 in Helsinki (University of Helsinki, Department of Dentistry and THL). The training was planned by the Health 2011 project team who consulted key experts in each special topic (see Appendix 1).

The general training of all fieldwork personnel covered an introduction to the aims and protocols of the study, ethical issues, data protection and informed consents, quality assurance, safety instructions, the roles and responsibilities of the main office and fieldwork personnel as well as rules and principles for communication. In addition, general IT training and an introduction to the working hours, travel arrangements, vacations and allowances were given.

The special training was tailored to the content of the respective measurement stations and it covered lectures and practices for interviewing techniques and measurements. The content and guidelines for each measurement station are explained in more detail under the specific chapters dealing with the contents of the health examination (see Chapters 5-19). As each nurse was appointed to work at two different examination stations, they received a special training for both of them. Before the actual fieldwork, the personnel had between one to two days to practice the study protocol with volunteers.

Training for the dentists and dental nurses was organized at the Department of Dentistry (see Chapter 13). Training of nurses working in the COURAGE sub-study is described in Chapter 19.2.

At the end of the training period, the fieldwork personnel were asked to fill in a feedback questionnaire. About 70 per cent of the personnel returned the questionnaire. Over 90 per cent of the fieldwork personnel stated that their training was at least partly sufficient. They also experienced that the atmosphere during the training was good and the training had increased their motivation and interest to begin the fieldwork. Although most nurses were quite satisfied with theoretical training, they would have needed more practical training on carrying out the measurements. Especially the training for home-visit nurses was considered insufficient and therefore some additional training was organized (see Chapter 2.5). Some fieldwork personnel also felt that they would have needed more IT training related to the study logistic system, the Blaise program and the Composite International Diagnostic Interview (CIDI mental health interview). In addition, they mentioned that some technical problems with the computer software, and the fact that some changes were made to the protocol during the training caused confusion.

2. IMPLEMENTATION

The fieldwork proper was conducted between 8 August and 21 December 2011. Complementary data collection continued until June 2012.

2.1 Fieldwork locations, facilities and personnel

Tomi Mäki-Opas, Ulla Laitinen and Annamari Lundqvist

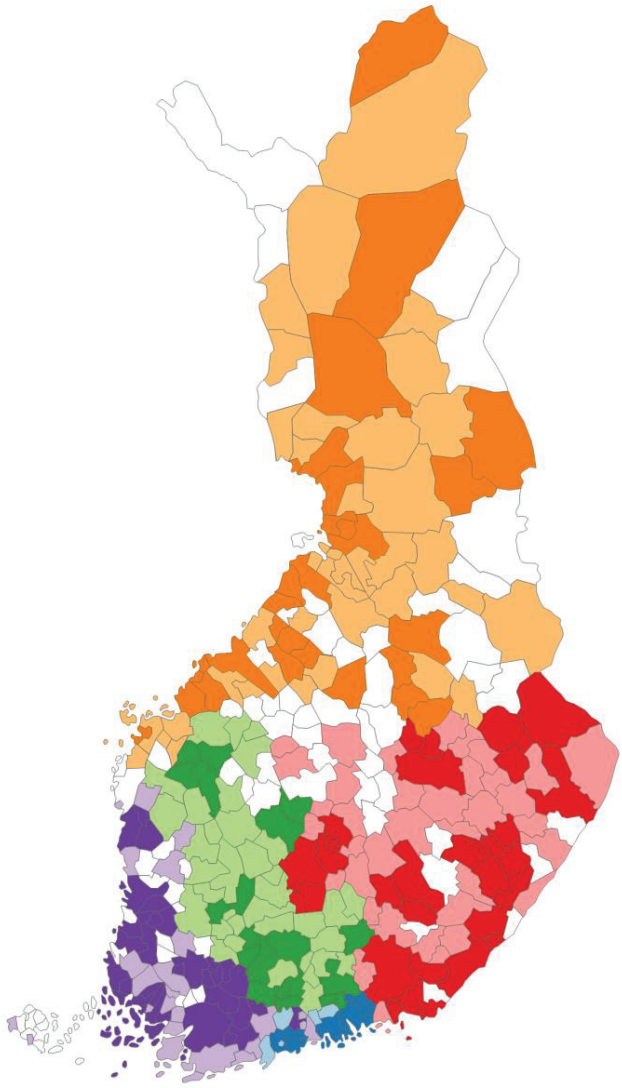
Fieldwork locations and facilities

A total of 60 municipalities around Finland were chosen as locations for the health examinations (see Figure 2.1.1). These locations were divided between the five fieldwork teams taking into account the number of subjects to be examined in each location, the number of days to be spent in each location as well as travel times and distances between locations. The number of locations varied between 3 (Helsinki team) and 16 (Oulu and Kuopio teams, see Appendix 3). Detailed timetables were drafted for each team. The teams travelled between locations in private and rented cars bringing along all required examination equipment. Accommodation was arranged at hotels, if needed.

The facilities required for the health examinations included 12 furnished examination rooms located in close proximity to one another and a separate corridor with a length of 18 meters for carrying out the walking test. The availability of facilities for the health examination was first asked by a letter sent to the chief physicians of the 60 municipal health centres. Many of the facilities were provided by the local health centres while some were rented from the private market.

The fieldwork was conducted by five fieldwork teams which consisted of 15 to 17 trained nurses (see Appendix 2 and Table 2.1.1). In Helsinki and Oulu, the team included also a dentist and a dental nurse.

Figure 2.1.1. Map of Finland including the study areas.



Different colours represent different university hospital regions. Darker shades correspond to actual Health 2000 health centre districts which where sampled and lighter shades correspond to areas where some of the members in the sample migrated between year 2000 and year 2011.

Table 2.1.1. The composition of field teams.

Nurse	Measurement station
Nurse 1	registration and final examination
Nurses 2 and 3	examination 1 and interview
Nurses 4 and 5	examination 2 and interview
Dentist	clinical oral examination (Helsinki, Oulu)
Nurse 6	clinical oral examination (Helsinki, Oulu)
Nurses 7, 8 and 9	examination 3 and interview
Nurses 10 and 11	examination 4 and Physical Activity and Fitness tests (sub-study)
Nurses 12 and 13	examination 4 and interview
Nurse 14	head nurse, final examination
Nurses 15 and 16	concise health examinations at home and substituting other team members when needed
Nurses 17, 18 (and 19)	home visits in the COURAGE sub-study

All nurses were trained to master work at two measurement stations of the health examination which ensured quality maintenance of measurements as well as flexibility and variation in work contents and schedules. This arrangement also enabled substitution during sickness absences. Moreover, two members of each fieldwork team were trained to conduct concise health examinations at home but they were also able to act as backup workers at the health examination sites during the busiest days or in case of sickness or other absences.

The head nurse in each team was in charge of the final measurement station and several other duties such as contacting the municipal health examination sites beforehand, keeping in touch with the main office, arranging accommodation for fieldwork personnel and transport between locations, answering queries of study subjects and co-ordinating the teamwork and implementation of the whole health examination protocol.

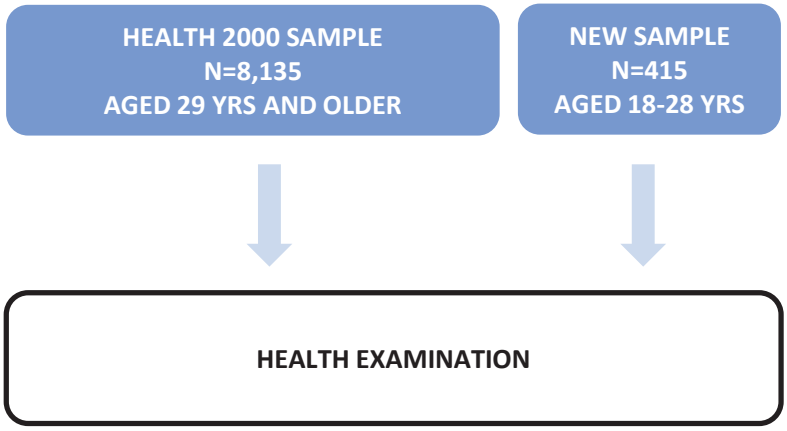
Between January and June 2012, the complementary data collection was conducted by six nurses working individually in the five main districts mentioned above.

2.2 Recruitment of the participants

Annamari Lundqvist, Päivikki Koponen and Seppo Koskinen

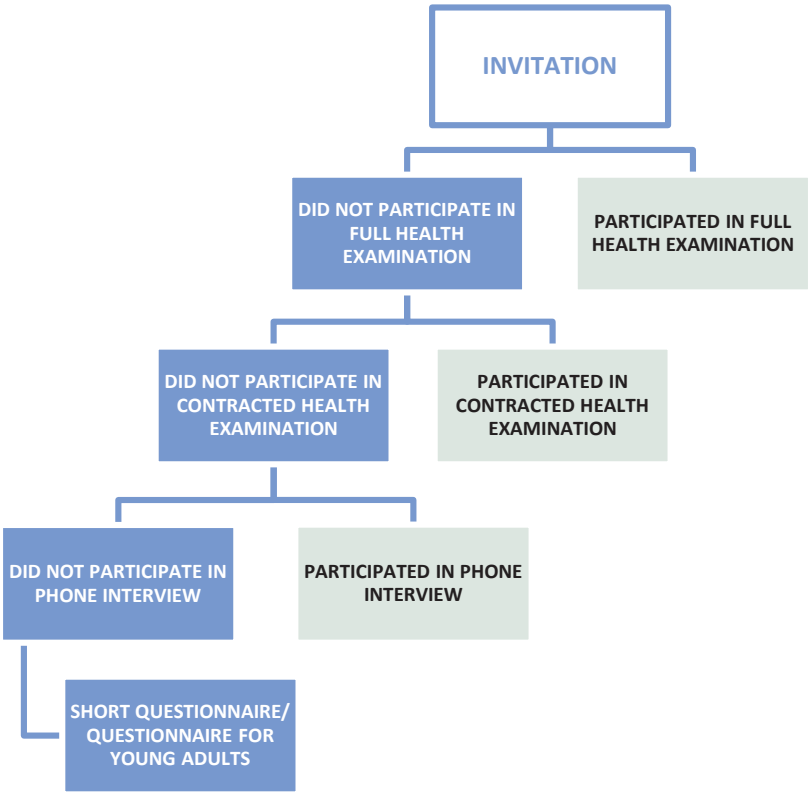
All subjects aged 29 years and older were invited to participate in the health examination (Figure 2.2.1). Young adults belonging to the sample of the COURAGE and the Physical Activity and Fitness sub-studies were also invited to the health examination.

Figure 2.2.1. Recruitment of the participants.



All persons unable or not willing to attend the full health examination were invited to participate in a concise health examination at home and, if this was not possible, to attend a phone interview (Figure 2.2.2). A short questionnaire (for subjects aged 29 years or older) or questionnaire for young adults (for subjects aged 18–28 years) was mailed to those who did not participate either in the health examination or the phone interview.

Figure 2.2.2. The recruitment process for those who did not participate in the full health examination in Health 2011 Survey.



2.3 Invitation and informed consent

Noora Ristiluoma, Annamari Lundqvist and Tomi Mäki-Opas

Invitation

Every person in the sample received an invitation letter approximately two weeks before the scheduled health examination. The invitation letter contained:

1) an information sheet which contained information on the objectives of the Health 2011 Survey, measurements included in the protocol as well as further use of survey and register data,

2) the appointment time for the health examination and the instructions on how to get to the examination site (attached map),

3) Questionnaire 1 (basic questionnaire, form T4002 for adults, T4002_2 for young adults),

4) a leaflet with information about the survey (form T4054), and

5) two copies of the informed consent form (form T4050).

Informed consent

At the beginning of the health examination, participants were informed about the Health 2011 Survey and they were offered a possibility to ask questions concerning the survey. After the nurse had made sure that the participants were aware of the purpose and conduct of the study, they were asked to sign two copies of the informed consent (form T4050). One copy was given to the participant and the other copy was collected for archival at THL. By signing the consent form, the participants gave permission that the information and samples gathered during the health examination, linked with other data (from registers), could be used for medical research.

If the participant was unable to fill in the consent form for reasons of poor health or limited cognitive capacity, the consent was signed by an accompanying family member or relative and a note "*on behalf of the participant*" was made on the form with the oral approval of the participant. The participants belonging to the DNA sub-sample were asked to sign a separate consent for taking and using DNA samples.

The nurse verified with his/her signature that all necessary information about the study had been given to the participant and that the participant's informed consent was received. The participants were informed that they have a right to withdraw from the study or cancel their informed consent later should they so wish.

2.4 Health examination

Tomi Mäki-Opas, Sebastián Peña and Annamari Lundqvist

The health examination protocols are presented in Figures 2.4.1, 2.4.2 and 2.4.3. More detailed information on the contents of the health examination is available in the Chapters 5-19.

On average, each fieldwork team was scheduled to carry out up to 24 health examinations starting between 8 AM and 3 PM at 15 minute intervals. On arrival and departure days, only 12 appointments were made as it took approximately 2 hours to set up and pack the examination points. In addition, some double appointments were occasionally needed to ensure a sufficient number of examination times. Spare times were also necessary for rescheduling appointments.

All subjects were asked to participate in seven different measurement stations within the health examination. In addition, subjects belonging to certain sub-study samples were asked to participate in additional measurements.

Each health examination was scheduled to last approximately 4 hours. The duration of the health examinations varied between 3.5 and 6 hours.

The health examination started with registration where the identity of the subject was first verified and the informed consent was obtained (see Chapter 2.3). Also the vision test was performed and Questionnaire 1 (which had been mailed to the subject attached with the invitation) was received and checked. Four questionnaires including Questionnaire 2 (form T4003 for adults and T4003_2 for young adults), Questionnaire 3 (form T4005 for adults and T4005_2 for young adults), a food frequency questionnaire (form T4006) and feedback questionnaire on experiences of the health examination (form T4023), were given to the participants for completion either during the examination or at home after the health examination. The content of the questionnaires is illustrated in Appendix 5.

In the first examination room, height was measured and body composition was analysed (bioelectrical impedance analysis). A 12-lead ECG measurement was also performed. A measurement of balance was carried out on subjects aged 70 years or older.

In the second examination room, blood pressure was measured twice from the right arm using a standard mercury manometer. In addition, heart rate and waist circumference were measured and a chair stand test was performed.

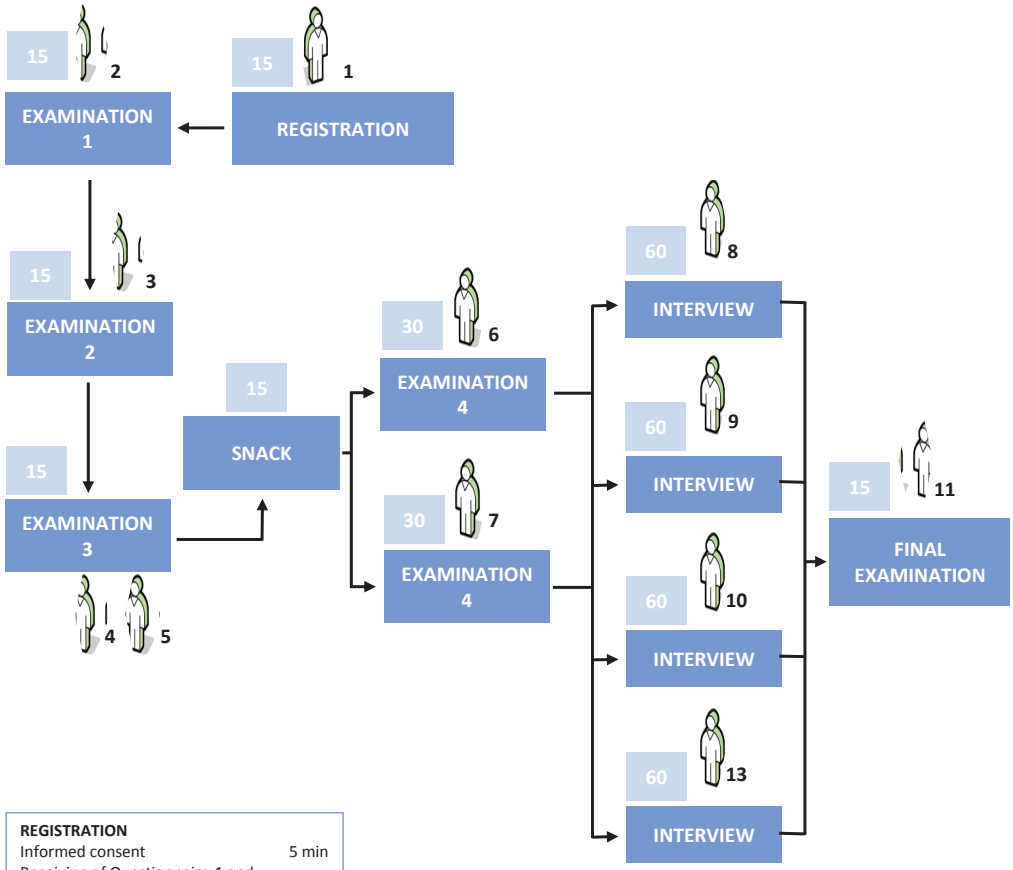
In the third examination room, fasting blood samples were taken and a spirometry test was administered. After that, participants were given a small snack and some time to fill in questionnaires given at the beginning of the examination.

In the fourth examination room, physical, cognitive and psychological functioning was assessed and the CIDI mental health interview was conducted.

Before the final examination station, subjects participated in an extensive health interview (form T4001 for adults and T4142 for young adults). It included questions about the subjects' background, living conditions, work and work ability, health behaviour, health and diseases, functioning, health service use and oral health (see Appendix 4).

In the final examination station, the completed questionnaires were returned. If participants had difficulties in answering questions, they were helped by the personnel. Joint functions tests were conducted for subjects aged 55 years or older. Participants were also given oral and written feedback on the results of measurements (see below), and if they had questions concerning the examination, these were answered. Questionnaire 4 (form T4045) was given along with a return envelope. As a final step the nurse confirmed that the subject had participated in all the examinations described above.

Figure 2.4.1. The health examination protocol including the number of fieldwork nurses per measurement station and estimated time (min) per measurement.



REGISTRATION

Informed consent	5 min
Receiving of Questionnaire 1 and handing Questionnaire 2 and 3	4 min
Handing FFQ and Feedback questionnaire	2 min
Visual acuity	3 min

EXAMINATION 1

Height	2 min
Bioimpedance	4 min
Resting ECG	8 min
Balance[70+]	2 min

EXAMINATION 2

Blood pressure (x2) and heart rate	10 min
Waist circumference	2 min
Chair stand test	2 min

EXAMINATION 3

Blood sampling	8 min
Spirometry	5-6 min

EXAMINATION 4

Cognitive functioning	9 min
Hand grip strength	4-5 min
CIDI mental health interview	15 min
Walking speed (6,1m)	1-2 min

[55+] Only for those aged 55 yrs and older
[70+] Only for those aged 70 yrs and older

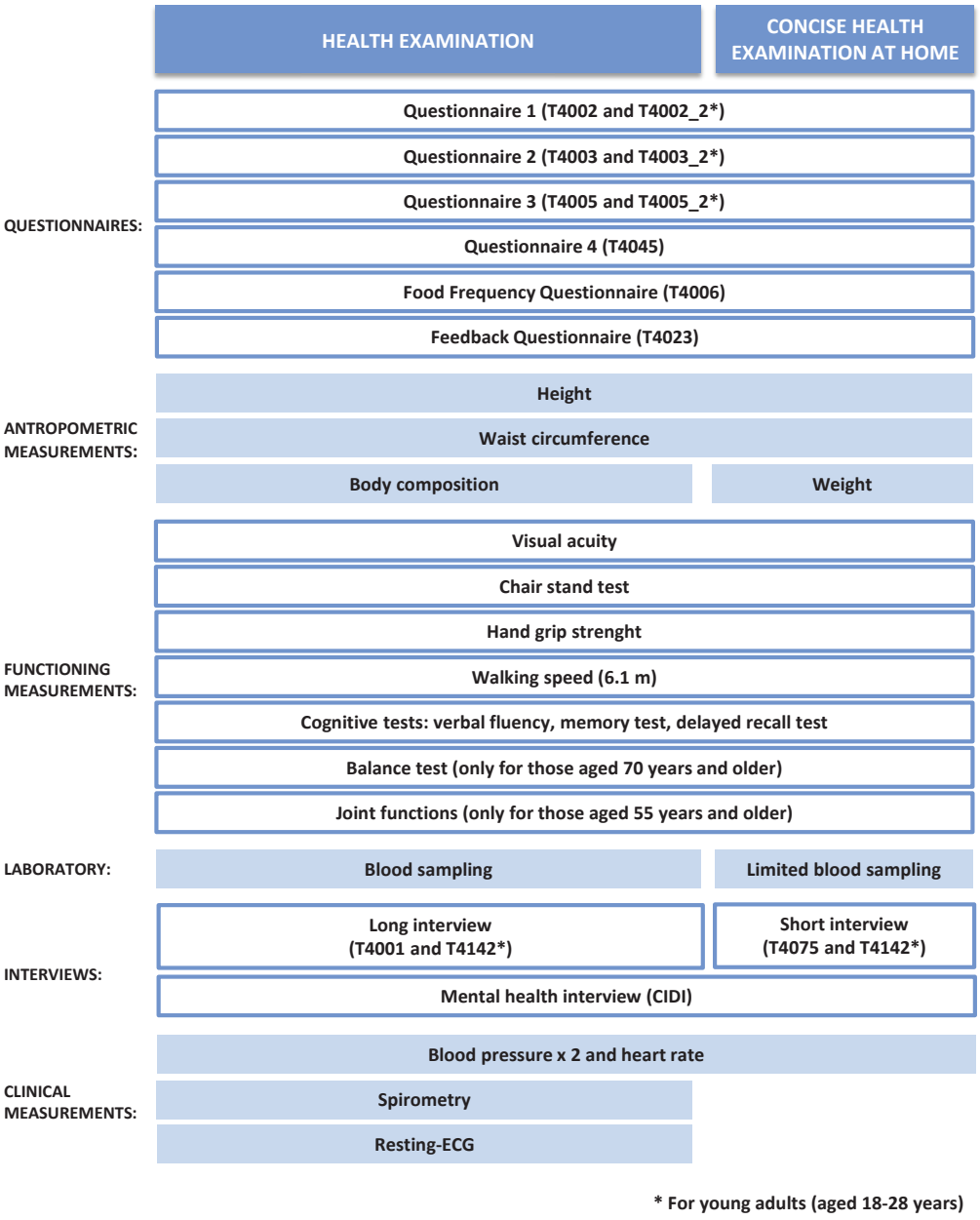
INTERVIEW

Background
Living environment
Health and diseases
Functioning
Health services
Health behaviors
Work and work ability
Oral health

FINAL EXAMINATION

Receiving Questionnaires	8 min
Joint functions [55+]	2 min
Handing Questionnaire 4	1 min

Figure 2.4.2. Content of health examinations.

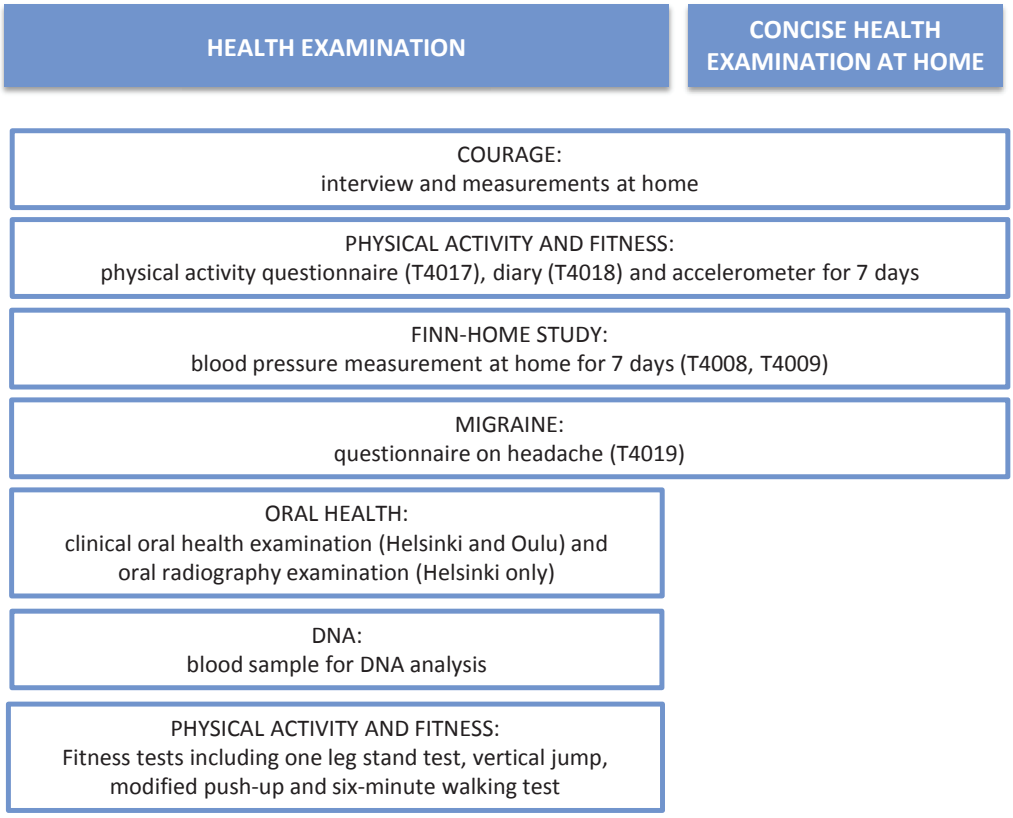


After the main phase of the data collection, between January and June 2012 (i.e. during the complementary data collection period), slightly abbreviated health examinations were conducted for the subjects who had not participated earlier. The slightly abbreviated health examination included all the above mentioned measurements except ECG. Furthermore, an abbreviated version of the interview (form T4075, see Appendix 4) was conducted.

Sub-studies in Health 2011 Survey

The Health 2011 Survey comprised several sub-studies illustrated in Figure 2.4.3 and in Chapter 19. For the subjects of the COURAGE sub-study, a home-visit comprising an interview and measurements was conducted generally about 1 or 2 weeks prior to the health examination. The clinical oral health examinations were carried out in two (Helsinki and Oulu) field examination areas, and the subjects in Helsinki were also invited to participate in the radiographic oral examination. Subjects reporting severe (moderate or intense) headaches in the health interview received an additional questionnaire on migraine. A blood sample for DNA analysis was drawn from those who did not have DNA available from their earlier participation in the Health 2000 Survey. At the end of the health examination, an automatic blood pressure measuring device together with instructions and a diary were given to the subjects belonging to the sub-study on blood pressure measurement at home. The subjects of the Physical Activity and Fitness sub-sample were directed to the fitness measurements conducted at the end of the health examination.

Figure 2.4.3. Content of sub-studies in Health 2011 Survey.



Feedback to the participants

At the end of the health examination, the participants were given written information on their height, weight, body mass index (BMI), waist circumference and body fat percentage as well as resting ECG, blood pressure and spirometry results (form T4022). In addition, the subjects received an assessment of their near and distance vision as well as the results of the tests measuring physical functioning comprising the chair stand test, the walking speed test and the hand grip strength test. On completion of the clinical oral examination, the participants received a written summary of the examination findings, a CD copy of the radiograph (only in Helsinki) and, when necessary, a recommendation to seek dental treatment.

A few months after the health examination, the participants received a letter (form 4130) containing the results of the following laboratory tests: total cholesterol, high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), triglycerides (TG), apolipoproteins A-l (ApoA1) and B (ApoB), gamma-glutamyltransferase (GGT), alanine aminotransferase (ALAT), C-reactive protein (CRP), glucose (Gluc), calcium (Ca), urate and creatinine. The letter also included an explanation of these results, the reference ranges and what to do if the results were not within the reference ranges. If the laboratory results differed significantly from the reference values, the participant was personally contacted by phone or (if this was not possible) by mail.

2.5 Concise health examination at home

Annamari Lundqvist

A concise health examination at home was performed on subjects unable to attend the health examination sites. Problems in physical or cognitive functioning were the main reason for home-based health examinations but also some subjects with small children, for example, preferred the concise health examination at home.

Concise health examinations were mainly conducted by home-visit nurses (two to three per each field team) during the main fieldwork period. In addition, six nurses carried out concise health examinations individually from January 2012 to June 2012. A three day additional training period was organized for home-visit nurses in August 2011 and those nurses who continued data collection after the actual fieldwork period were further trained in January 2012 (2.–5.1. and 9.–12.1.). The training covered the measurements and other tasks for the home visits as well as a supervised concise health examination at home.

The concise health examination protocol mostly comprised the same measurements as the full health examination (Figure 2.4.2). The measurements of blood pressure, heart rate, height and waist circumference were similar to those in the full health examination. Weight was measured with a portable scale and at home there were fewer blood samples due to difficulties of proper processing. Tests on functional ability were mostly similar but some adaptations had to be made due to the circumstances at home. The CIDI mental health interview was carried out for all participants at home while the health interview was abbreviated to save time (form T4075, see Appendix 4).

2.6 Phone interview and short questionnaire

Annamari Lundqvist

Phone interview

Phone interviews were offered to subjects not willing or unable to participate in the health examination or the concise health examination at home. The average duration of the phone interview was 15–20 minutes. Both fieldwork nurses and members of the project team working at the main office conducted the phone interviews.

The phone interview (form T4077) consisted of an abbreviated version of the original interview and selected questions from Questionnaire 1. It included questions about the subjects' background, health, illnesses, health service use, oral health, health behaviour, functioning and work ability (see Appendix 4). At the end of the phone interview, subjects were again offered a possibility to participate in the health examination or concise health examination at home, and some of them seized this opportunity.

Short questionnaire

The short questionnaire with a return envelope was mailed to all subjects who had not been reached or who were not willing or able to participate in the health examination or the phone interview. The questionnaire was also mailed to non-respondents who had not totally refused to participate in the survey.

The content of the short questionnaire (form T4095) was almost identical with the content of the phone interview, comprising selected key questions from the interview and Questionnaire 1 (see Appendices 4 and 5). It included questions about the subject's background, health, illnesses, health service use, oral health, health behaviour, functioning and work ability.

2.7 Questionnaire for young adults

Annamari Lundqvist

The majority (N=1,579 persons) of the sample of young adults (aged 18–28 years) received the mailed self-administered questionnaire for young adults only. The questionnaire (form T4140) included questions about the subject's background, health, illnesses, health behaviour, work, psychological well-being and quality of life (see Appendix 5). A paper form with a return envelope was sent to the subjects but they also had the possibility to fill in an internet-based questionnaire. All subjects returning the questionnaire received a small gift (value of 8 euros) and took also part in a lottery (value of main prize about 500 euros). Up to two reminders (one letter including the questionnaire and one postcard) were sent to non-respondents.

The questionnaire for young adults was also mailed to those young adults who were invited in health examination but did not participate in the health examination proper or the phone interview.

2.8 IT environment and logistics system

Mikko Pekkarinen, Marko Grönholm, Talvikki Leinonen and Esa Virtala

Equipment and software

Local area networks and hardware used by fieldwork teams included:

- Laptop (DELL Latitude E6420) encrypted HDD (Utimaco)
- Encrypted flash drives for backups (Kingston DataTraveller Locker)
- Cable lock (TARGUS DEFCON CL)
- Internet stick (Nokia CS-17 or Nokia Sierra)

Each nurse had a computer with internet access via an internet stick (DNA broadband). An e-mail box was created for each nurse on THL e-mail server so that they could easily communicate with the project main office via Office Outlook Web Access (OWA). In addition, the fieldwork teams could use a web browser to read and print out guides and forms from the

Health 2011 website. Some examination stations (e.g. those with ECG, spirometry or bioimpedance) had printers (HP LJ 1102w, HP Office Pro 8000 and Samsung SCX-3205), barcode scanners (Datalogic, Quickscan M2 130 BLK) and other hardware.

In the first laptop, the Microsoft Windows XP Professional operating system as well as all the software and drivers required in examination stations were installed. The rest of the laptops were then cloned by THL's information management unit. Data collection was mainly conducted by electronic forms created with the Blaise system, a computer-assisted personal interview (CAPI) system developed by Statistics Netherlands. Blaise version 4.8 was used in the electronic data capture at various health examination points.

Every field team had also mobile phones to communicate with the main office and with the study subjects.

Training

IT training of the fieldwork personnel highlighted the accuracy of the data recorded and data security. This training consisted of two parts:

- All fieldworkers:
  - a brief introduction to the use of data recording equipment in general
  - an instruction on the use of the software at their respective examination stations
- Two persons from each fieldwork team:
  - to set up and operate the local network of the fieldwork site

Most of the IT support and guidance were provided by phone from THL main office. Field visits were also made. The data collection programs were usually updated via the Health 2011 extranet pages or by sending an encrypted memory stick to one of the specially trained members of the fieldwork team. In some cases, the software was uploaded from the Health 2011 operations manager´s workstation via file transfer protocol (FTP) to the field group's workstations.

It was important that the Health 2011 IT operations manager at THL was able to link up directly from his/her own workstation to the field teams´ workstations. This was particularly

useful for carrying out certain checks and in troubleshooting and repairing problems. Support for the use of data entry software was provided by the people who had written the software.

Logistic system

The logistic system was primarily used for the following functions:

- THL appointments system
- Invitation lists and field teams' calendars
- Stock accounting of samples and paper forms
- Recording of completed visits
- Communication between THL and the study subjects (invitation letters, laboratory results etc.)

The main database was Oracle Rdb relational database version 7.2 on an Open VMS system. Programming of the system was primarily based on XSQL pages and XSL transformations producing HTML pages. Various listings and reports were produced as HTML pages by using this technique. The pages could be accessed and printed out using standard browsers both at project main office and in the field office. The system was deployed in Apache Tomcat application server and the requests were directed through Apache HTTP server, which was the front-end server to Tomcat.

Transfer files were analysed automatically by special Perl and Blaise Manipula programs for collecting inventory data as well as information on specimens taken at examinations. These data were stored into the study database for quality control purposes.

2.9 Communications

Tomi Mäki-Opas and Eeva-Liisa Kallonen

The aim of communications in the Health 2011 Survey was to gain positive publicity for the project. Immediately at the beginning of the fieldwork, both local and nationwide media were informed about the Health 2011 Survey in order to raise general awareness and to encourage active participation.

Communications team

Communication was organized centrally at THL. Tasks of the communications team included nationwide informing at the beginning of the survey.

As soon as the fieldwork started, the fieldwork teams took more responsibility for regional communications. The head nurse in each fieldwork team was in charge of contacting the local radio stations and newspapers to increase the local visibility of the survey. They also acted as liaisons between the field and the main office.

Channels and publicity

National communications was carried out both via nationwide and regional newspapers as well as television and radio news. Before the start of the fieldwork, a press release was issued by the main office. The media were also encouraged to visit the examination sites. TESSO magazine – a magazine of THL and Ministry of Social Affairs and Health – published a report of the Health 2011 pilot.

The personnel of THL were familiarised with the survey through an information session organised at THL. Moreover, in the electronic journal for the personnel, there was a section devoted to the study.

A leaflet with information about the survey was sent to the study subjects in the invitation letter.

The Health 2011 Survey website acted as an important channel of communication. The website was opened in spring 2011 and contained information about the background, objectives and contents of the survey.

Overall, the project succeeded well in communications by achieving widespread visibility and publicity in both national and regional media. According to media monitoring, publicity was highest (over 40 media results) in August 2011 when the fieldwork started. News coverage was generally positive or neutral.

3. PARTICIPATION

Annamari Lundqvist, Harri Rissanen, Esa Virtala, Tommi Härkänen, Sebastián Peña and Paul Knekt

3.1 Response rates

Of all those invited to the Health 2011 Survey (i.e. Health 2000 sample and new sample of young adults), 67.3 per cent (N=6,740) participated in at least one phase of the data collection (i.e. health examination, phone interview or questionnaire, Figures 3.1.1 and 3.1.2 and Tables 3.1.1 and 3.1.2).

Health 2000 sample

Among the living members of the original Health 2000 sample who resided in Finland and were invited to the Health 2011 Survey, the participation rate was 73.6 per cent (N=5,903, Figure 3.1.1 and Table 3.1.1). The participation rate was 69.9 per cent among men and 76.8 per cent among women.

Altogether 59.0 per cent (N=4,729) of the sample participated in the health examination. This included 4,277 subjects participating in the full health examination and 452 subjects participating in the concise health examination at home. Phone interviews were conducted on 5.0 per cent (N=404) of the sample and the short questionnaire was returned by 9.6 per cent (N=770) of the subjects.

A total of 5,784 subjects participated in both the Health 2000 and the Health 2011 Surveys. There were 1,734 subjects participating in Health 2000 Survey only, 119 subjects participating in Health 2011 Survey only and 385 subjects participating neither in 2000 nor in 2011.

Figure 3.1.1. Participation in the Health 2011 Survey (Health 2000 sample).

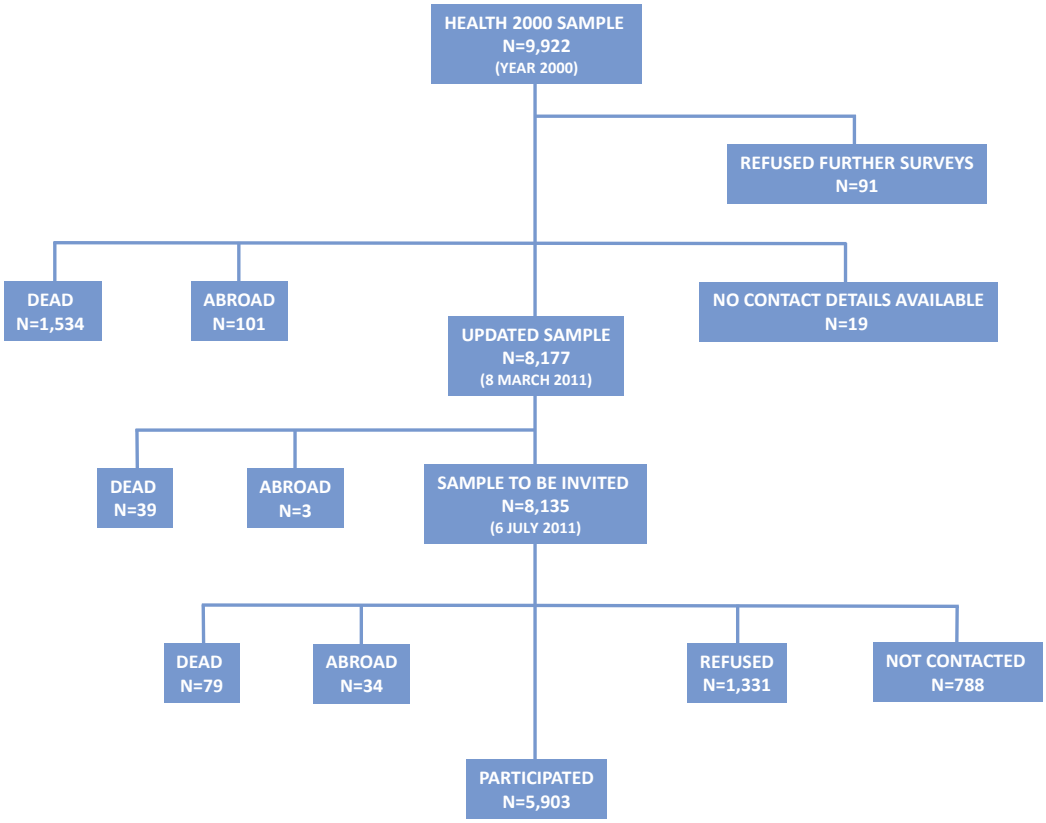


Table 3.1.1. Participation in different stages of data collection by sex and age (Health 2000 sample).

	Final sample	Health examination		Phone interview		Short questionnaire		At least one	
	n	n	%	n	%	n	%	n	%
Men (yrs)									
29–40	934	312	33,4	63	6,7	128	13,7	503	53,9
41–50	816	463	56,7	41	5,0	72	8,8	576	70,6
51–60	816	490	60,0	26	3,2	83	10,2	599	73,4
61–70	698	487	69,8	29	4,2	44	6,3	560	80,2
71–80	377	263	69,8	15	4,0	22	5,8	300	79,6
81–	138	92	66,7	4	2,9	9	6,5	105	76,1
All	3,779	2,107	55,8	178	4,7	358	9,5	2,643	69,9
Women (yrs)									
29–40	866	411	47,5	64	7,4	128	14,8	603	69,6
41–50	873	579	66,3	35	4,0	88	10,1	702	80,4
51–60	898	594	66,1	57	6,3	80	8,9	731	81,4
61–70	744	544	73,1	33	4,4	52	7,0	629	84,5
71–80	500	328	65,6	25	5,0	40	8,0	393	78,6
81–	362	166	45,9	12	3,3	24	6,6	202	55,8
All	4,243	2,622	61,8	226	5,3	412	9,7	3,260	76,8
All (yrs)									
29–40	1,800	723	40,2	127	7,1	256	14,2	1,106	61,4
41–50	1,689	1,042	61,7	76	4,5	160	9,5	1,278	75,7
51–60	1,714	1,084	63,2	83	4,8	163	9,5	1,330	77,6
61–70	1,442	1,031	71,5	62	4,3	96	6,7	1,189	82,5
71–80	877	591	67,4	40	4,6	62	7,1	693	79,0
81–	500	258	51,6	16	3,2	33	6,6	307	61,4
All	8,022	4,729	59,0	404	5,0	770	9,6	5,903	73,6

New sample of young adults

Among the new sample of young adults invited either to the health examination or to fill in the self-administered questionnaire, the participation rate was 42.3 per cent (N=837, Figure 3.1.2

and Table 3.1.2). Of those invited to the health examination, 29.8 per cent (N=121) took part in the health examination and 22.9 per cent (N=93) participated in the phone interview or returned the self-administered questionnaire. Altogether 39.6 per cent (N=623) of those who received the mailed self-administered questionnaire for young adults took part in the survey.

Figure 3.1.2. Participation in the Health 2011 Survey among young adults (new sample, aged 18–28 years).

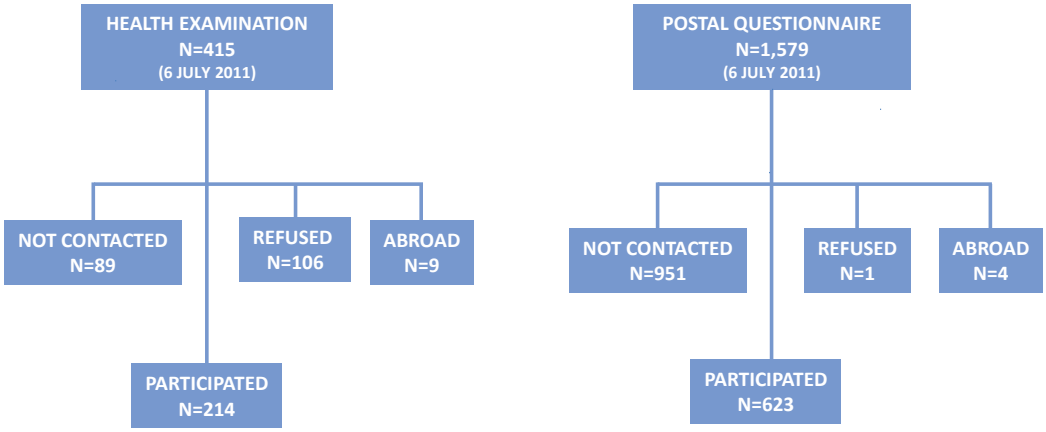


Table 3.1.2. Participation in different stages of data collection in the Health 2011 Survey among young adults (new sample, aged 18–28 years).

	Final sample	Health examination		Phone interview		Mailed questionnaire		At least one	
	n	n	%	n	%	n	%	n	%
Health examination	406	121	29.8	24	5.9	69	17.0	214	52.7
Questionnaire	1,575	-		-		623	39.6	623	39.6
Both samples (total)	1,981	121		24		692		837	42.3

Mini-Finland resurvey

The participation rate among those who had participated in the Mini-Finland Survey in 1978–1980 was 81.3 per cent (N=748). There were 723 subjects participating in all three surveys i.e. in 1978–1980, 2000–2001 and 2011.

Participation in sub-studies

Participation in sub-studies is illustrated in Table 3.1.3.

Table 3.1.3. Participation in sub-studies in the Health 2011 Survey.

	Final sample (n)[1]	Participated (n)	Participation rate (%)
Physical Activity and Fitness	4,821	2,455	50.9
COURAGE	3,858	1,972	51.1
DNA	3,135	886	28.3
Finn-Home Study	1,836	1,194	65.0
Oral health examination	3,878	1,845	47.6
Oral radiographic examination	2,208	1,023	46.3
Migraine	875	832	95.1

[1)] Includes samples of Health 2000, Mini-Finland and young adults.

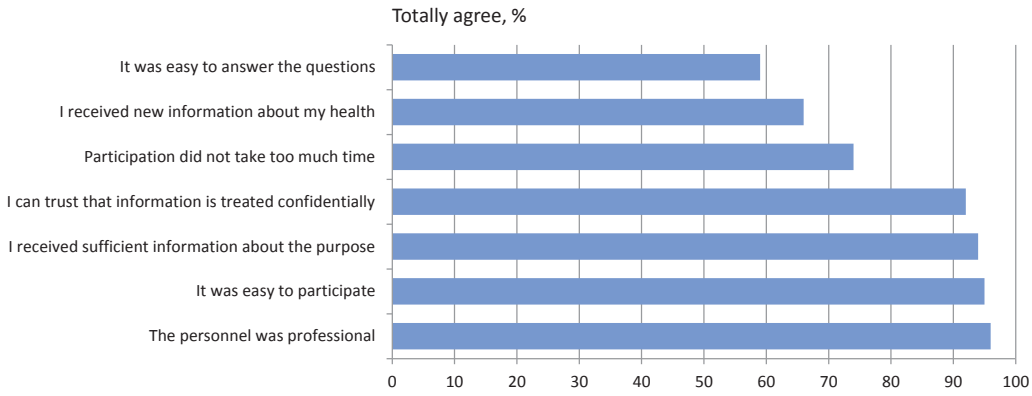
3.2 Experiences of the participants

Tuija Jääskeläinen and Päivikki Koponen

The participants were asked to fill in a feedback questionnaire after the health examination or to complete within two weeks an electronic questionnaire on the Health 2011 website. Altogether 44 per cent of men and 45 per cent of women who participated in the health examination provided feedback.

Feedback received from the respondents was mainly positive (Figure 3.2.1). The participants experienced that the fieldwork personnel were professional and behaved politely and were friendly (as many as 99 per cent totally agreed). The majority also felt that they received sufficient information about the objectives of the survey and that it was easy to participate in the health examination. Furthermore, the respondents could trust that their information was treated confidentially. One respondent summarized: "*The health examination was very accurate and comprehensive. The staff members were friendly and truly professional, thank you"*. Around one fourth of the respondents, however, felt that participation took too much time. One-third of men (33 per cent) and almost half (46 per cent) of women also had some difficulties in answering the questions/questionnaires.

Figure 3.2.1. Feedback from the participants (per cent of respondents).



The respondents were mainly very satisfied (over 90 per cent completely satisfied) with all measurements and the way how different parts of the health examination were carried out. The only exception was the snack; 77 per cent of the respondents were completely satisfied with the snack offered to them (coffee or tea, juice, sandwich, fruit and cookies). However, one respondent pointed out: *"The snack was too small compared to the time we had to be fasting (without eating any other food)"*.

3.3 Efforts to increase the participation rate and reasons for non-participation

Annamari Lundqvist, Päivikki Koponen and Tuija Jääskeläinen

Several efforts were made to increase the participation rate.

- If the preliminary appointment time for health examination was unsuitable, the participants had a possibility to reschedule their appointment time by contacting a toll-free number at the main office in THL.
- On the day before the examination, a reminder was sent by SMS to those participants whose phone number was available.
- If the subjects did not arrive to health examination, they were contacted as soon as possible to arrange a more suitable appointment time. Three attempts at different weekdays and times of the day were made to contact the non-attendants by phone.
- Invitations to take part in the survey were sent by SMS to those subjects who could not be reached otherwise.
- If the home address was available, the fieldwork nurses tried to reach subjects who had not been contacted by phone or SMS by visiting their home.
- Every effort was made to remove and overcome obstacles that could prevent subjects from taking part in the health examination. For example, people with physical disabilities were offered free taxi rides to and from the health examination sites. If necessary, the use of interpreters was also utilized when possible.
- Phone numbers were available for about two-thirds of the sample. Some of them were, however, found to be wrong or not in use anymore. Subjects with prepaid phones or secret numbers were not listed in telephone directories. A large amount of work was done to search phone numbers e.g. in the internet.
- The self-administered short questionnaire or the questionnaire for young adults was mailed to all those who were invited but did not participate in the health examination or phone interview. Up to two reminders were mailed to non-respondents. All subjects who filled in the questionnaire received a small gift (value of 8 euros) and took also part in a lottery (value of prize about 500 euros).
- To gain positive publicity for the project and to encourage active participation, both local and nationwide media were informed about the Health 2011 Survey.

The short questionnaire included questions about the reasons for refusing to take part in the health examination. The main reasons for refusal were that the given appointment time or place

were not suitable. Almost half (40 per cent) of the respondents reported that they would participate in future if they could choose the appointment time and place. The respondents also wished that they would receive more information about their own health. In fact, almost one-third of the respondents reported that they would have participated if the survey had included a clinical medical examination conducted by a physician. The fact that a medical examination was included in the survey in 2000 but not in 2011, may partly explain why quite many of those who participated in the Health 2000 Survey did not want to participate again in 2011. Moreover, in the Health 2000 Survey the appointment for the health examination was booked personally during the interview while in the Health 2011 Survey a pre-arranged appointment time was given in the invitation letter. Even though a toll-free number was given to change the appointment, making the phone call may not have been easy enough for everybody. Another important difference in comparison to the Health 2000 Survey was the lack of a personal contact between the study subjects and the field work personnel to find a suitable personal appointment time. Nevertheless, one-fifth of those who reported reasons for non-participation indicated that they would not have participated in any situation.

4. DATA

4.1 Collection, protection and management of the data

Pirkko Alha, Tuija Jääskeläinen, Helmi Koskinen, Mikko Pekkarinen, Sirkka Rinne, Harri Rissanen, Noora Ristiluoma and Esa Virtala

Data collection process

Based on experiences from the Health 2000 Survey, it was decided that as much data as possible would be collected electronically, mainly using the Blaise program. Personal data were stored in the logistic system and information from the measurements of bioimpedance and spirometry were directly stored into an electronic format by the measurement devices. For each participant, a separate file was also produced for the CIDI mental health interview. Similarly, information on ECG measurement was stored in a separate file.

In each health examination station, the data were first recorded in the workstation's encrypted hard disk. On a daily basis, stored data were saved to the flash drive and transferred in batches to the main office at THL. Sometimes files were transferred less frequently than once a day due to limited internet access. Furthermore, ECG data were stored on memory cards which were sent to THL in Turku for further processing. Spirometry, bioimpedance and CIDI mental health data were sent from THL to the collaborators for further processing. The data collected using paper forms were sent to an outside company for data entry.

All data collected and recorded were saved as SAS files using the version 9.22 of SAS software for statistical analysis (SAS Institute Inc 2010).

Evaluation of the data collection process

Perhaps the biggest IT problems during the study were caused by late planning and preparation for key aspects of the data collection process; indeed parts of the IT processes and programs were finalised only during the training and some updating was needed during the first days of the fieldwork. The fact that the necessary hardware equipment and programs were fully operational only once training for fieldwork personnel was underway complicated the training process.

Overall the targets set for project IT operations were reasonably well met. Problems were encountered during the course of the project, but they were all relatively minor and manageable. The hardware concept was reliable and easy to use. The fieldwork personnel quickly learned the saving routines, and preferred the electronic form of recording the data. Standards of data security were high.

The appointments system and other parts of the logistics system were sound and useful tools for organizing the health examinations and the collaboration between the fieldwork teams and the main office. The existence of completed forms and electronic data could easily be checked and followed, especially as the forms were barcoded. Barcodes and appropriate programs also helped to ensure speedy and accurate recording of the sample tubes. The same solutions and procedures have been applied in other THL surveys, too.

Data protection and data security

Data protection was taken care of at all stages of the Health 2011 Survey. During the fieldwork, each participant received an identification number (ID code) to be used in bar code stickers attached to all paper forms and samples (blood tubes), and all personal data were kept in locked boxes. The same ID codes were used in electronic files. To ensure privacy and data protection, data collected in the Health 2011 Survey are treated as confidential under the obligation to professional secrecy.

Safe storage of the electronic data collected in the health examination was ensured by several different ways. First, the laptops used by the field teams had encrypted hard disks (Utimaco SafeGuard Easy, version 3.20); even if a laptop would have gone missing or had been stolen, the data would have remained safe. All workstations had a theft-prevention sticker

(SAFEREG) and they were protected by F-Secure antivirus software and F-Secure Firewall. Communications between the main office and the field teams took place over DNA broadband and the connection was secured with Cisco AnyConnect VPN. Nonetheless, all electronic materials dispatched by the field teams were encrypted (*"secure copy"*), as was all HTTPS communication from main office to the field teams (so-called https protocol was used).

All hardware at the main office in THL is protected by a firewall, and access is restricted to authorized users only by usernames and passwords. Access to the health examination material and questionnaire data is restricted to only a few members of authorized THL personnel. Specific study ID codes instead of the person's name or contact information are used in all Health 2011 datasets and thus, the researchers analysing the data do not have access to any data where persons can be identified.

Personal information is, however, needed for follow-up purposes and for linking with other data including national registers. It is therefore recorded in a separate dataset table which can be accessed by a small number of persons only. When the Health 2011 dataset is linked with other data, special attention is given to data protection. This is ensured by close adherence to the relevant legislation, the rules of THL and the bodies maintaining the registers as well as the guidelines of good research practice.

Checking of the data

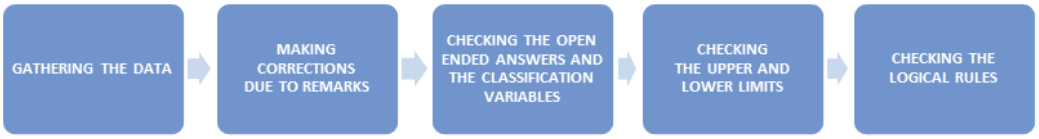
Most of the data were checked and corrected at THL while data from bioimpedance measurements, spirometry and ECG as well as the CIDI mental health interview were checked by external collaborators. Data checking required a great deal of work and time, partly due to limited testing of data collection methods before the fieldwork was started.

Checking of the data included several steps illustrated in Figure 4.1.1. As a rule, these checks were carried out only on data from the same examination station or form, i.e. no checks were carried out between different forms.

In regard of interview data, a first a dataset was created in which every respondent had only one record. Thus, for example, the information from the short and phone interviews was included into the long interview for respondents who had not finished the interview at the examination site but instead the interview had been completed at home or by phone.

For the questionnaires, no preliminary checks were carried out but the forms were sent directly to an outside company for data entry. In general, the quality of the recorded data was found to be acceptable. However, due to the complex structure of the dietary questionnaire (FFQ), the data entry had to be done twice.

Figure 4.1.1. The steps of checking and correcting the data.



The Blaise program used in the health examination and the interview had many built-in checks and thus, acceptable values and the logic of the series of questions were checked at data entry. The study nurses conducting the health examinations and interviews could enter any relevant notes and comments in the Blaise program's REMARK field. During the data checking, remarks were used to complete missing information.

Some questions ended with the response option *"other, what?"*. If the respondent had chosen this option, the first step was to check whether any of the response alternatives provided was an appropriate answer. If so, the corresponding changes were made to the variables. In some cases new variables were created.

The interviews included questions such as those related to medications and illnesses that had preset classifications. These classifications had been incorporated into the Blaise program to facilitate the interviewer's job. The interviewer first entered the medicine or illness reported by the interviewee as text and then chose the corresponding code from a menu that appeared on the screen. For example, ATC codes were used for medicines. The self-reported illnesses were coded in the same way as in the Health 2000 data by following the classification originally used by KELA. The classification includes the ICD-9 and ICD-10 disease codes under the headings of different disease categories – altogether close to one hundred classes.

The upper and lower limit checks were mainly used for questions with open-ended numerical answers.

The main focus in the logical checks was on the logical consistency of the responses to the main question and follow-up questions. Priority was given to the follow-up question. For instance, if the respondent gave the number of cigarettes smoked, the response to the main question (*"Do you smoke?"*) was corrected to *"yes"* in cases where the recorded response was "*no"*.

Response criterion was specified for each examination station or form. In general, at least two variables with acceptable data had to be available.

The variables based on identical questions in the long interview and a shorter form (e.g. phone interview) were labelled according to the variable tag used in the long interview. Corresponding variables from the Health 2000 Survey were given the same names in the Health 2011 data. The variables that differed from those in the Health 2000 Survey were distinguished with the abbreviation T11.

Several widely used scores for instruments such as General Health Questionnaire, Maslach Burnout Inventory and 15D quality of life measures were created. Also some single variables had to be re-coded to make them comparable with the Health 2000 variables.

SAS software (SAS Institute Inc 2010) was the main tool used in data processing (checks, creation of derived variables, etc.).

4.2 Quality assurance and evaluation of the fieldwork

Päivikki Koponen, Paul Knekt, Harri Rissanen, Tuija Jääskeläinen, Markku Heliövaara and Arpo Aromaa

Fieldwork quality assurance

During the fieldwork proper, the experts in charge of the measurements, interviews and examination stations audited the field operations during their visits to the fieldwork sites. In addition to helping resolve any problems, these visits were an important source of encouragement and also evaluation tool. Each visitor stayed at the survey site for one or two days, during which actual measurements were observed with the consent of the participant.

There were also discussions with the fieldwork team leader and the nurses. Feedback was given directly after the observations or later by e-mail.

During the audit visits the following issues were observed:

- How is the communication between team members and participants going?
- How is the informed consent obtained?
- How is the privacy of the participants ensured?
- How is the data security ensured?
- How are the measurement devices used?
- How often and how are the instruments checked and calibrated?
- How are the measurements performed?

Personnel from the main office discussed also with the team leader about:

- Selection of the survey premises
- Resources and time schedule
- Recruitment of the participants
- Data management
- Contacts with the local media
- Difficulties met and achievements

From the site visit, a written report was prepared and distributed among the main office personnel, if feasible.

The head nurses had three one day meetings in Helsinki to discuss problems and experiences during the fieldwork, and to report back on the progress made in the field. At the same time they were informed about current and upcoming events to be shared with the rest of the fieldwork team.

Experiences of the personnel

Feedback from the fieldwork personnel was collected with a questionnaire at the end of the fieldwork. A total of 69 per cent of the personnel returned the questionnaire.

The fieldwork personnel experienced that co-operation with the main office was quite smooth, but some of them reported that they would have expected more feedback and support from the main office during the fieldwork. Furthermore, the majority of the personnel were satisfied

with the coordination and management of the survey while technical problems, including problems with the computer software especially at the beginning of the study, were reported to have caused most difficulties. The personnel were, however, satisfied with the IT support. One of them reported: *"The members of the IT support team were helpful and easy to reach, there were no problems with them"*.

Almost all of the fieldwork personnel rated the accommodation during the fieldwork at least acceptable. Accommodation was mainly arranged in local hotels. In some localities, the fieldwork personnel had to stay overnight at the health examination site, and they felt that they did not have adequate possibilities to rest and relax. The fieldwork personnel also reported some disagreement about the timetables related to travelling from one locality to another. While many members of the field teams were pleased with the flexible working times, some mentioned that there was more work at evenings than expected. In addition, some nurses were dissatisfied with the salaries, e.g. since they did not get any extra payment for working outside regular office hours, while they received such additional payments when employed in health centres and hospitals.

The majority of the personnel were satisfied with the accuracy of the timetables of fieldwork and co-operation within their field team. They were also mainly pleased with the planning and coordination of the fieldwork by the head nurse but about one-fifth of them reported that they were disappointed with these arrangements within their own field team. Due to the difficulties in recruiting fieldwork personnel, the level of expertise within the teams and the role of the head nurses varied, as not all of them had experience of health surveys, fieldwork or in any similar team leadership position. In the previous Health 2000 Survey, the head nurses were recruited earlier than the other team members, so that they were better familiar with the survey protocols before the fieldwork started and had more resources for team leadership. This would have been beneficial in the Health 2011 Survey, but the head nurses could not be recruited earlier due to the limited budget.

The fieldwork personnel were very satisfied with the way the participants were treated. One of them mentioned: *"I experienced that every nurse of our field team treated participants well and properly"*. The personnel agreed with the participants that the study was comprehensive and included diverse measurements. However, they had also noticed that some participants expected that there would have been a same kind of a clinical medical examination as in the Health 2000 Survey. Some nurses felt also that the participants, especially older subjects, were expected to fill in too many questionnaires and they had difficulties in answering some questions. Moreover, reaching by phone and motivating those participants who did not come to

the health examination was considered partly frustrating when nurses were instructed to contact the non-participants many times and at the end in some cases without any results. On the other hand, some nurses reported that by contacting the participants they succeeded in motivating them to participate.

Quality control

Quality control in the Health 2011 Survey aimed to obtain high quality data that were comparable between fieldwork teams and observers. The aim was also to ensure comparability between the Health 2000 and the Health 2011 Surveys, as far as possible, in order to allow analysis of time trends both at individual and population levels.

Quality control measures were taken to monitor the survey process, so that any problems could be detected at an early stage, and actions were taken to correct the detected problems as soon as possible.

A large number of quality control measures were built into the Health 2011 Survey at different stages of the survey process. The following issues were taken into account:

Prior to fieldwork:

- Sampling to ensure survey representativeness, and obtaining register data to evaluate and minimize the effect of non-response (see Chapter 1.2)
- Piloting the survey instruments and the feasibility of the study protocol (see Chapter 1.4)
- Documented survey procedures to ensure standardized stable measurements. These were available for all fieldwork personnel at the survey website, and are included in this report (see Chapter 2 and Chapters 5-19)
- Training of the fieldwork personnel to use the standard procedures during the fieldwork (see Chapter 1.4)

During fieldwork:

- Computer-assisted data entry with predefined upper and lower limits of most values to prevent typing errors (see Chapter 4.1)
- Calibration and other quality control for the measurement devices (described in the

chapters for each measurement)

- Audit visits by the trainers and other experts from the main office
- Rotation of personnel between measurement points to avoid bias caused by the observers
- Data management: due to the shortage of personnel at the main office, some errors in the data could have been avoided by more frequent and systematic data checking during the fieldwork, avoiding time to be spent later for data checking and corrections (see Chapter 4.1)
- Repeated measurements (see results presented in quality analyses below)

After fieldwork:

- Data management, checking and correction (see Chapter 4.1)
- Statistical analysis to evaluate the overall repeatability of selected measurements (see results presented in quality analyses below)

Quality control was carried out by the survey team at the main office, i.e. as internal quality control. Due to restricted resources it was not possible to have the survey observed and assessed by an independent outside body (external quality assessment), except for the field laboratory activities (see Chapter 5). However, as the principal investigator, the project leader, and most of the survey personnel at the main office were well experienced, and top experts in the field, the need for assessment by an independent outside body was not considered necessary.

Quality assessment

The quality of the Health 2011 Survey data has been assessed on the basis of pre-defined criteria. The survey data was evaluated and documented at the main office. For different data items different issues were checked. These included:

For questionnaires/interviews:

- comparability of questionnaires, i.e. how well the Health 2011 question(s) corresponded to the questions used in the previous Health 2000 Survey and whether there were differences in data collection methods, i.e. self-administered questionnaire vs. interview

- data extraction, i.e. how well similar key indicators could be extracted from the new data (compared to the previous Health 2000 data)
- proportion of missing data for each questionnaire/interview item
- distribution of all categorical variables
- minimum values, maximum values, 10% percentile, median, 90% percentile and standard deviation for continuous variables
- logical checks to identify values outside specified ranges

For measurements:

- recording form/Blaise comparability with the previous Health 2000 recording form/Blaise
- measurement protocol and details in the new 2011 manuals in comparison to the year 2000 manual
- differences in instruments (new instruments were used in some measurements, e.g. bioimpedance, hand grip strength and spirometry)
- proportion of missing data
- minimum values, maximum values, 10% percentile, median, 90% percentile and standard deviation
- for blood pressure: proportion of identical readings between subsequent measurements and distribution of terminal digits
- for anthropometric measurements: proportion of various terminal digits

Quality analyses

During the fieldwork the following repeated or parallel measurements were carried out:

- At the end of the health examination the respondents received Questionnaire 4 with selected questions from Questionnaire 1 and the interview. The participants returned Questionnaire 4 later by mail. In the quality analyses the successive answers to the same questions were compared.
- During a few days in each fieldwork team (weekly or whenever feasible), a few participants were asked to take part in repeated measurements. If they agreed, the home visit nurses repeated some measurements carried out earlier by other nurses.

The overall repeatability of the responses to the questionnaire/interview items given during an average of 13 days interval (median 5 days) and the agreement between the parallel measurements taken within 2-3 hours were estimated using the weighted kappa coefficient (Fleiss 1981) in case of a polytomous or dichotomous variable and using the reliability coefficient (Winer 1971) in case of a continuous variable. The significance of differences between the first and second assessment was tested with the symmetry test (for polytomous measures), McNemar's test (for dichotomous measures) and the paired t-test (for continuous measures). The statistical analyses were carried out using SAS software, version 9.22. (SAS Institute Inc 2010)

The repeatability of questions concerning hindrance caused by back, neck, shoulder or joint pain was relatively high (Table 4.2.1). The reliability coefficient was 0.81 for hindrance at free time and 0.83 for hindrance at work. Both measures showed statistically significantly lower values at the second assessment occasion.

Table 4.2.1. Repeatability of selected continuous variables.

Question	N	α	1st mean (SD)	2nd mean (SD)	p	Variable
On a scale from 1 to 10, please give an estimate on how much hindrance is caused by your back, neck, shoulder or joint pain:						
Hindrance at work	4,420	0.83	2.70 (2.58)	2.65 (2.57)	0.03	OIRE_65A
Hindrance during free time	4,685	0.81	2.75 (2.47)	2.71 (2.50)	0.06	OIRE_65B

N = number of observations; α = reliability coefficient; p = p for difference; SD = standard deviation

With one exception, the overall repeatability of the questions concerning health status was relatively good, the kappa coefficients varying within a relatively narrow interval (0.73-0.83, Table 4.2.2). The only exception was the series of questions concerning ability to perform daily tasks (kappa 0.62). Of the nine questions, three showed a statistically higher prevalence at the first and three at the second measurement occasion.

Table 4.2.2. Repeatability of selected categorical variables.

Question	N	K	95% CI	1st prev.	2nd prev.	Variable
Do you have any permanent or chronic illness or any defect, trouble or injury?	4,722	0.77	0.75-0.79	55.7	52.2	Kys1_K04
Have you ever had hay fever (seasonal allergic rhinitis) or another type of allergic rhinitis?	4,841	0.83	0.81-0.85	39.0	39.4	OIRE_29
Have you had back pain…[1)]	4,827	0.76	0.74-0.77	38.5	43.6	OIRE_46A
Have you had neck pain…[1)]	4,812	0.74	0.72-0.76	34.6	39.2	OIRE_52A
Have you had shoulder pain during…[1)]	4,805	0.77	0.75-0.79	27.1	32.7	OIRE_56A
Have you had pain, ache or motion sensitivity in one or more joints…[1]	4,743	0.80	0.79-0.82	58.5	56.6	OIRE_57
Have you been unable to perform your daily tasks or duties due to back, neck, shoulder or joint pain…[1)]	4,590	0.62	0.58-0.66	9.8	8.9	T11_Kys2_K47
Have you had difficulty in walking or have you limped due to a hip disorder or defect…[1)]	4,754	0.73	0.71-0.76	17.8	17.4	OIRE_59
Have you had difficulty in walking or have you limped due to a knee disorder or defect…[1)]	4,730	0.76	0.74-0.79	21.6	21.5	OIRE_60
How much do you exercise and strain yourself physically in your leisure time?	4,761	0.71	0.69-0.73	nr	nr	Kys1_K27
How often do you have a drink containing alcohol?	4,732	0.89	0.88-0.90	nr	nr	T11_Kys1_K28
How would you rate your quality of life?	4,817	0.68	0.66-0.70	nr	nr	T11_Kys1_K41
Do you ever feel lonely?	4,890	0.75	0.73-0.76	nr	nr	T11_Kys1_K52
How would you describe the current balance between income and expenditure in your household?	4,884	0.80	0.78-0.82	nr	nr	Kys1_K09

N = number of observations; K = Kappa statistics; 95% CI = 95% confidence interval; 1st prev. = first prevalence; 2nd prev. = second prevalence; nr = not relevant

[1)] …during the last 30 days?

Of the lifestyle factors the repeatability of the question concerning alcohol consumption was high (kappa 0.89) whereas that concerning exercise was lower (kappa 0.68). Both variables showed a statistically significant change ($P < 0.001$) between the two assessment occasions. The repeatability of the questions concerning quality of life varied from 0.68 to 0.80. Also these measures showed a statistically significant change.

The overall agreement for the prevalence of asthma, hypertension, diabetes, and mental disorders was excellent (kappa 0.89-0.96) and that for general health, osteoarthritis, and oral health was good (kappa 0.73-0.79, Table 4.2.3). Also the determination of physical activity was reliable (kappa 0.76-0.86) whereas it was only satisfactory for memory (kappa 0.68). Of the two lifestyle variables considered, the agreement for smoking was excellent (kappa 0.96) and for consumption of fruit and vegetables was good (kappa 0.85). The agreement of the functions of joints was relatively good for some measures (kappa 0.70-0.75) but unsatisfactory for some others (kappa 0.30-0.41). Generally no notable systematic differences between the measurement occasions were found.

In general, the overall agreement between observers was good for the continuous variables (Table 4.2.4). The reliability coefficients for waist circumference, maximal hand grip strength, and work ability varied from 0.94 to 0.97 and for all other variables between 0.63 and 0.81. Only for systolic blood pressure, times for stand ups, and walking speed, the two observers reported systematically different results. Based on the reliability coefficients the quality of the blood pressure and heart rate measurements may be seen as satisfactory. However, it is important to bear in mind, with regard to blood pressure and heart rate measurements, for example, that repeatability is affected not only by factors that have to do with the actual measurement process, but also by biological variation in the subject.

Table 4.2.3. Repeatability of selected variables related with health, illness and functioning.

Question	N	K	95% CI	1st prev.	2nd prev.	Variable
Present state of health good or rather good	118	0.74	0.61-0.88	70.3	75.4	BA01
Asthma diagnosed by a doctor	118	0.96	0.87-1.00	10.2	11.0	BA04
High blood pressure, hypertension diagnosed by a doctor	118	0.96	0.91-1.00	33.9	33.9	BA13
Osteoarthritis (arthrosis, joint degeneration) diagnosed by a doctor	114	0.73	0.57-0.89	22.8	17.5	BA19
Psychological or mental illness diagnosed by a doctor	118	0.89	0.76-1.00	13.6	12.7	BA25
Diabetes diagnosed by a doctor	118	0.90	0.76-1.00	9.3	9.3	BA26
Condition of teeth and the health of mouth good or rather good	118	0.79	0.66-0.92	73.7	79.7	EA01
Manage shopping without difficulties	114	0.76	0.49-1.00	94.7	93.9	HA02H
Able to walk about 2 kilometers without resting without difficulties	118	0.86	0.75-0.98	81.4	80.5	HB08
Very good or good memory (self-reported)	117	0.68	0.54-0.81	58.1	65.0	HF10
Eating fruit or berries daily	118	0.85	0.75-0.94	56.8	54.2	FA10
Daily smoking	118	0.96	0.89-1.00	13.6	12.7	FB05
Normal squatting	65	0.60	0.37-0.83	73.9	81.5	NIVEL_1
Normal abduction of the right upper arm	65	0.55	0.11-0.99	92.3	96.9	NIVEL_2O
Normal abduction of the left upper arm	65	0.65	0.20-1.00	95.4	95.4	NIVEL_2V
Normal external rotation of the right shoulder joint	65	0.38	0.09-0.67	80.0	87.7	NIVEL_3O
Normal external rotation of the left shoulder joint	65	0.41	0.11-0.71	83.1	86.2	NIVEL_3V
Normal internal rotation of the right shoulder joint	65	0.75	0.60-0.91	40.0	49.2	NIVEL_4O
Normal internal rotation of the left shoulder joint	65	0.70	0.51-0.89	69.2	72.3	NIVEL_4V

N = number of observations; K = Kappa statistics; 95% CI = 95% confidence interval; 1st prev. = first prevalence; 2nd prev. = second prevalence

Table 4.2.4. Repeatability of selected measurements.

Question	N	α	1st mean (SD)	2nd mean (SD)	p	Variable
Systolic blood pressure (mmHg)	115	0.71	136.4 (17.5)	138.7 (17.5)	0.07	MIT1_SYSTBP
Diastolic blood pressure (mmHg)	115	0.67	81.3 (10.1)	81.0 (10.0)	0.68	MIT1_DIASTBP
Heart rate	115	0.63	67.9 (10.8)	67.5 (11.2)	0.62	MIT1_HRATE
Waist circumference (cm)	116	0.96	93.5 (12.9)	93.7 (12.8)	0.47	MIT1_CIRCUMWAIST
Hand grip strength	118	0.94	38.3 (12.8)	38.2 (12.6)	0.86	PUR_MAX
Work ability (scale from 1 to 10)	108	0.97	7.98 (1.84)	7.95 (1.90)	0.55	IB04
Chair stand test (five times, s)	114	0.77	11.3 (3.6)	10.3 (3.6)	<0.001	TUO_3_2
Chair stand test (ten times, s)	114	0.71	23.8 (7.5)	21.8 (9.5)	0.001	TUO_4_2
Walking speed (m/s)	118	0.81	1.84 (0.39)	1.89 (0.45)	0.04	KAV_NOPEUS2

N = number of observations; α = reliability coefficient; p = p for difference; SD = standard deviation

The quality control of bioimpedance and spirometry measurements is described in Chapters 6.3 and 7.3, respectively. Information on quality assurance of oral health examination is available in Chapter 12.

Quality control in blood sampling

The field laboratory personnel were trained in advance, and the sampling and sample handling sites were audited once by an external auditor. Also the person responsible for the laboratory work audited the work in each fieldwork team sites once or twice. Quality control in the field laboratory has been described in more detail in Chapter 5.

4.3 Data from registers

Päivikki Koponen, Seppo Koskinen and Harri Rissanen

Data obtained for the sample from the National Population Register comprised information on some key characteristics of each person (Table 4.3.1). In addition, administrative register data were obtained with specific permissions sought from the institutes/organisations responsible for each register. Register linkages were made using the personal registration number assigned to all residents in Finland. Register data are used for several purposes e.g. to analyse non-response, and to obtain additional information on sociodemographic characteristics, health status and on the use of health services and social security benefits before, during and after the survey. Register data can also be used to study whether the use of social security and health services is adequately based on the needs identified in the survey data.

A comparison of the participants' and non-participants' health in the light of register based information also helps to evaluate the accuracy of the results obtained by questionnaires and health examination. Register data were used to assess the characteristics of the non-participants and to construct the survey weights to be used in the analysis (see Chapter 4.5). Register based follow-up provides incidence data enabling future epidemiologic studies to find out how the survey data predict the development of the participants' health, by linking the cross-sectional survey data with follow-up data on the participants' causes of death and illnesses, as well as use of services and social insurance benefits/allowances.

The record-linkage was designed and carried out in close co-operation between the project organisation and the bodies maintaining the registers concerned. The most important register data are those on causes of death, hospital treatments, entitlements to medication reimbursements and certain other illness-related benefits, purchases of prescribed medicines, cancers, work disability and employment as well as housing.

Table 4.3.1. Register data available for the Health 2000 and Health 2011 samples.

Organisation	Topics
Population Register Centre	Age, sex, date and place of birth, marital status, place and type of residence, mother tongue
Ministry of Employment and the Economy	Employment service register data: periods of unemployment and participation in labour market training and work/training trials
The Social Insurance Institution (KELA)	Coverage by the social insurance in Finland Disability, rehabilitation and sickness allowances Reimbursement for medicine expenses Purchases of selected medicines Allowances for pensioners
National Institute for Health and Welfare (THL)	Care Register for Health Care (inpatient care and outpatient visits, diagnoses and operations and other care procedures) Register of Primary Health Care Visits Cancer Registry (diagnosed cancers) Mass Screening Registry (mammography and pap smear) Medical Birth Register (year(s) of giving birth, prenatal care and care during births) Register of Induced Abortions (year(s) and types of procedures) National Infectious Diseases Register (selected diagnosed diseases) Register of Social assistance (household receiving social assistance)
Statistics Finland	Education, occupation and socioeconomic status Causes of death
Finnish Centre for Pensions	Earnings related pensions

4.4 Using the data for research purposes

Seppo Koskinen, Harri Rissanen and Pirkko Alha

The data of the Health 2011 Survey are available for research purposes in collaboration with the project organisation. In order to obtain access to data, researchers must first submit a study proposal which will be reviewed by the Health 2011 Scientific Board.

The forms to apply access to the data are available on the website of the Health 2011 Survey. The health 2011 website includes all the forms used during the fieldwork and the

corresponding variables. A code number, given to each approved study proposal, is required to order the data.

An agreement of co-operation is required. One signed copy of the agreement is given to the researchers and the other one is archived in THL.

The sample collection of the Health 2011 Survey has been transferred to THL Biobank in June 2015.

4.5 Statistical analyses

Tommi Härkänen and Paul Knekt

The description of the sampling design and missing data analyses as well as an empirical comparison of different statistical methods to handle missing data in the Health 2011 Survey have been described elsewhere (Härkänen et al. 2016). There is more documentation and some examples to analyse the Health 2011 Survey data in the internet (http://www.thl.fi/terveys2011). Here an overview of the most important aspects to account for in practical statistical analyses, and the background of the weights needed to account for the sampling probabilities and non-response is given.

Important notes on statistical analyses

The longitudinal Health 2011 Survey data allow researchers to estimate changes both in the cross-sectional population distributions and longitudinal changes on an individual level. More reliable information on possible causal relations between the observed quantities can be obtained using representative population survey dataset consisting of repeated measures and register-based follow-up linked with them than by using cross-sectional data possibly linked with register follow-up which can, however, be easier to analyse.

The increased non-response requires attention, and researchers should assess the possible mechanisms, which cause non-response, and in each analysis perform corrective measures among which commonly applied methods are:

- **Weighting** of observations corrects the distribution of known background factors (age, sex etc.) in the group of participants to match the distribution in the population. Weighting methods are useful in correcting unit non-response. The weight variable should match the analysis variables, e.g. if the analysis involves variables collected in the health examination then the health examination weight should be selected. There are several weights, which are based on participation in different parts of the survey:
    - participation in the health examination
    - participation in any part of the survey (good for questionnaire variables)
    - participation in a sub-studies
- **Multiple imputation** is a more advanced and efficient method which can handle item non-response better than weighting, but it requires more experience in conducting analyses. In order to minimize bias, the imputation models must be constructed separately for each research problem. For the Health 2011 Survey variables it is advisable to incorporate appropriate Health 2000 Survey variables in the imputation model as the participation rate was much higher in 2000 than in 2011. E.g. risk factors, which can predict the outcomes of interest in 2011, are likely to be important.
- Non-response can depend (directly) on the variables of interest, e.g. healthy individuals participate and individuals with disease do not. This kind of **missing not at random (MNAR) non-response** can be very difficult or impossible to correct for by using any statistical methods without additional information such as register data on health status.

The original Health 2000 Survey was based on a complex sampling design, and the resurvey in 2011 increases the number of clustering levels by one. This should be accounted for by using proper statistical methods such as mixed-effects models, although the design-based methods, which have generally been used to analyse the cross-sectional Health 2000 Survey data, appear to provide similar results.

- The sample is **not a simple random sample** from the population, thus standard statistical methods assuming independence of observations do not generally produce correct estimates.
- **Geographical representativeness** is limited to continental Finland and the five university hospital districts as well as the biggest cities. In some provinces there is only one – usually the biggest – health centre district in the sample, and the results on such small areas are therefore not representative.

The Mini-Finland re-survey provides up to three measurements on the same individual sampling units, who were 30 years old or older in 1978–1980, and have been followed up to 33 years. The complication in analysing these data is that the study subjects were selected

geographically in the vicinity of the university hospital regions in year 2000. Therefore this sample does not represent the Finnish population. In the study by Stenholm et al. (2012) the effects of the selection mechanism were mitigated by analysing only the individuals who already at the time of the Mini-Finland baseline study lived in the same municipalities where the Mini-Finland resurvey was conducted.

The analyses on the Health 2011 Survey data can be conducted using most general-purpose statistical software packages. Multiple imputation can be conducted using, for example, SAS (SAS Institute Inc 2010), Stata (StataCorp 2009) or R (R Core Team 2012) software packages. Design-based as well as model-based mixed-effects analyses can be performed using Sudaan (Research Triangle Institute 2008), SAS, Stata or R (survey package, Lumley 2004; lme4 package, Bates et al. 2012) software packages. Model adjusted estimates based on the predictive margins can be calculated using Stata and Sudaan software packages.

Sampling probabilities of the new sample of young adults

The EPSEM sample of the Health 2000 survey yielded sampling probabilities, which were constant in the age groups of 18 to 79 years, and 80 years or older (Laiho et al. 2008). The EPSEM does not hold after year 2000, because the population sizes have changed. The true population sizes in year 2011 as well as in 2000 based on the boundaries in 2000 were obtained from Statistics Finland. The notation needed to calculate the sampling weights is presented in Table 4.5.1.

The inclusion probability of individual sampling unit i, who belonged to stratum $s := s(i)$ and health centre district $k := k(i)$, in year 2000 was written as

$$p_i^{00} := \frac{n_{sk}^{00}}{N_{sk}^{00}} \frac{m_s N_{sk}^{00}}{N_s^{00}} = \frac{m_s n_{sk}^{00}}{N^{00}} \quad \text{(Equation 4.5.1)}$$

The "size" of cluster k was the corresponding population size N_{sk}^{00} of age 18 or older. The sampling weight was defined as $v_i^{00} := 1/p_i^{00}$. Equation 4.5.1 reduced nicely, but in 2011, however, the PPS probabilities were the same as in 2000, thus the inclusion probabilities in 2011 did not reduce similarly. In other words, the sampling design was not self-weighting in 2011.

$$p_i^{11} := \frac{n_{sk}^{11}}{N_{sk}^{11}} \frac{m_s N_{sk}^{00}}{N_s^{00}} \quad \text{(Equation 4.5.2)}$$

The sampling weight was defined as above: $v_i^{11} := 1/p_i^{11}$.

Table 4.5.1. Notation corresponding to the population and sample sizes, and to the sampling weights. For the 15 largest towns, define $m_s := N_{s1}^{00} := N_{s1}^{11} := 1$.

	Year 2000	Year 2011
	AGED 18– YRS	AGED 18–28 YRS
Sample size	n_{00}	n_{11}
Number of strata	S	S
Population size in stratum s	N_s^{00}	N_s^{11}
Number of health centre districts sampled in stratum s	m_s	m_s
Population size in stratum s and health centre districts k	N_{sk}^{00}	N_{sk}^{11}
	ALL AGES	ALL AGES
Participation status	R_i^{00}	R_i^{11}
Right-censoring status		Z_i
Sampling probability	p_i^{00}	p_i^{11}
Expansion weight	w_i^{00}	w_i^{11}

Let $v_{s\,sx}^{x} := \sum_i v_{six}$ denote the sum of weights in stratum s and year $x \in \{00, 11\}$. The sums of weights v_s^{00} defined by Equation 4.5.1 were proportional to the population sizes of the strata in 2000, but not in 2011. Therefore we rescaled the weights in 2011 by

$$w_i^{11} := \frac{v_i^{11} N_s^{11}}{\sum_{i \in s} v_i^{11}}. \quad \text{(Equation 4.5.3)}$$

The new sample was drawn from the same areas as in the Health 2000 Survey, thus the original sampling probability p_k^{00} of cluster k was the same in 2011.

Adjusting the weights for non-response in cross-sectional studies

The studies which report results based on the Health 2000 survey use the baseline weights of year 2000 (Laiho et al. 2008). Weighted statistics such as means and prevalence provide representative results on the target population. In addition to the cross-sectional statistics, follow-up data can also be analysed using the baseline weights and the methods for cohort studies. The means and prevalences in the population in 2011 cannot be estimated using the baseline weights because of the (increased) non-participation in 2011 (Table 4.5.2).

First, define some notation. Let $1 / p_i^{00}$ denote the baseline sampling weight of individual sampling unit i. Let $R_i^{11} := 1$ for the participants of the Health 2011 Survey and $R_i^{11} := 0$ for others. The participation status in the Health 2000 Survey R_i^{00} was defined similarly by the so-called "respondent's union" criterion (Laiho et al. 2008). If individual sampling unit i died or emigrated between years 2000 and 2011, then let $Z_i := 0$, otherwise $Z_i := 1$. Let X_i denote baseline measurements and possible register-based follow-up data, which were associated with the participation R_i^{11} in year 2011. There was item non-response in the components of X_i. The missing values were single imputed using the `transcan` function of the `Hmisc` package (Harrell Jr and others. 2014).

Table 4.5.2. Participation in the Health 2000 and Health 2011 Surveys.

Participated in 2000, R_i^{00}	Participated in 2011, R_i^{11} Yes	No	Sum
Yes	5,602	1,589	7,191
No	301	643	944
Sum	5,903	2,232	8,135

The Health 2011 Survey weights for the participants in 2011 were constructed by dividing the original sample into two parts. The first part consisted of the participants of 2000 ($R_i^{00} = 1$).

The observed baseline information on these subjects is very useful, because the baseline information contains information on various risk and lifestyle factors predicting future diseases or functional disabilities, which are common causes of non-response. The sampling weights based on the sampling probabilities, which equalled the inverse of the sampling intervals, were updated based on that information and the *inverse probability weighting* method (Robins et al. 1994).

Every individual sampling unit alive and living in Finland in 2011 ($Z_i = 1$) made a decision to participate or not, and the participation ($R_i^{11} = 1$) probability was modeled here using a logistic regression model. The inverse of the probability is the non-response weight.

$$v_i^{11} := \frac{1}{\mathbb{P}_{\beta^{00}}\{R_i^{11}=1|X_i^{00},Z_i=1,R_i^{00}=1\}} = 1 + \exp\{-X_i^{00}\beta^{00}\}. \quad \text{(Equation 4.5.4)}$$

For $Z_i = 0$ we defined $v_i^{11} := 0$. We assumed that the non-response can be explained by the observed baseline or register-based variables X_i^{00} in Equation 4.5.4, that is, the missing data mechanism was assumed to be *missing-at-random* (MAR, Rubin 1987, Molenberghs and Kenward 2007). In reality, the missing data mechanism was likely to be *not-missing-at-random* (NMAR), in which case the non-response was likely to depend on unobserved factors possibly including outcome variables of analyses such as the health status at the time of the Health 2011 Survey.

The second part comprised the small group of non-participants of the Health 2000 Survey ($R_i^{00} = 0$). This allowed us to utilize the second group in the weighted analyses as well, but the covariate vector X_i^{00} only contained age group, sex, education and university hospital district.

The expansion weights were obtained by calibrating the product of the baseline sampling weights $1 / p_i^{00}$ and non-response weights v_i^{11} by age group $AG_i \in \{(18\text{-}28), (29\text{-}49), (50\text{-}74), (75\text{-}79), (80+)\}$ and participation in 2000 ($R_i^{00} \in (0,1)$) group $\mathcal{I}_{AG_i,R_i^{00}}$. The population sizes (excluding the immigrants after year 2000) in each group were approximated by $N_{AG_i,R_i^{00}}^{11} := \sum_{\ell \in \mathcal{I}_{AG_i,R_i^{00}}} Z_\ell / p_\ell^{00}$. The final, calibrated weight was then

$$w_i^{11} := v_i^{11} / p_i^{00} \frac{N_{AG_i,R_i^{00}}^{11}}{\sum_{\ell \in \mathcal{I}_{AG_i,R_i^{00}}} v_i^{11} / p_\ell^{00}}. \quad \text{(Equation 4.5.5)}$$

Selection of the weighting model

In the case of the new sample of young adults and in cross-sectional studies the only information available on non-respondents is usually register data, which contain only sparse information, such as age, sex, area and education. These register-based variables were therefore used as covariates in model (see Equation 4.5.3).

In the case of the original Health 2000 Survey the participation rates have decreased during the follow-up, which is typical in longitudinal studies. The baseline measurements, however, contained a wide variety of information on the participants of the Health 2000 Survey, and this information can be utilized to build a more realistic model on the non-response of the baseline participants in the follow-up study. Note that the original unit non-response at the baseline has been taken care of using the post-stratification weights, which will be updated using Equation 4.5.5.

Table 4.5.3. Wald tests of the non-response model, in which all these variables were entered into the model simultaneously.

Names	LR Chisq	Df	Pr(>Chisq)
Age group 10 years	89.32	5	0.0000
Sex	9.50	1	0.0021
Education	34.15	2	0.0000
Self-reported work ability	40.49	2	0.0000
Language (1=Finnish, 2=Swedish)	28.97	1	0.0000
Self-rated health status (10-class VAS)	2.73	1	0.0985
Time use: clubs or associations	96.11	4	0.0000
Age group 10 years : Sex	17.77	5	0.0032
Age group 10 years : Education	22.61	10	0.0123
Sex : Education	5.14	2	0.0764
Age group 10 years : Sex : Education	18.26	10	0.0507

In this work we selected numerous variables, which cover various areas of the information collected in the Health 2000 Survey. These variables were then entered into logistic regression models either univariately as main effects or in interaction with age and/or sex. Also the

interaction of age, sex and education was included in all models. The Bayesian information criterion (BIC, Schwarz 1978) was then applied to assess which variables had better predictive power in the non-response than the model containing only the interaction of education, age group and sex. The best predictors were then entered into the same model as main effects. Using the Wald test, the variables, for which the p-value was below 0.20, were selected into the final model (Table 4.5.3).

Joining the new sample and the original sample

In longitudinal analyses the data are usually arranged in the so-called "long" format, in which there is one row for each observation, in this case two rows for individuals, who participated in both surveys, and one row for individuals, who participated in one survey only. In order to conduct analyses on the joint data of the new sample of young adults and the original Health 2000 Survey sample, the expansion weights defined by Equation 4.5.5 are used for the Health 2011 Survey observations. The original weights of the Health 2000 survey are used for the baseline observations.

Methods to handle effects of missing data

Various methods have been proposed to handle effects of missing data (Molenberghs and Kenward 2007). In addition to the IPW and post-stratification (Lehtonen and Pahkinen 2004) methods described above, there are improved methods based on weighting (e.g. the doubly robust methods, Wirth et al. 2010), and other methods based on augmenting the missing data values.

Generally weighting can be appropriate in cases where the proportion of item-non-response is low. In other cases the missing data values in the few variables can be imputed, and all information contained in the partially observed sampling units can be utilized.

We use the baseline weights based on the design and participation in 2000. Their updated versions called Health 2011 Survey weights described earlier in this chapter are used for 2011. The baseline weights do not account for the increased non-response in 2011.

In multiple imputation (MI) the missing data values are imputed using a predictive distribution, which is based on the observed data and possible prior information (Rubin 1987, Schafer 1999). This imputation model can differ from the analysis model, which is applied on the imputed data containing no missing values. Generally there is considerable uncertainty in the imputed values, thus a single imputation would underestimate the uncertainty (variance) of the results. Therefore in MI several copies of the original dataset are created, both the imputation procedure and the statistical analyses are performed separately on each of them, and finally the results based on the imputed datasets are combined.

In a typical item-non-response case the analysis variables cannot be ordered to form a monotonic missing data pattern. This restricts the range of adequate MI methods. In standard statistical software packages the variables, which contain missing values, are assumed to follow a multi-normal distribution, which does not suit well to categorical or other non-Gaussian variables. Binary variables are, however, often approximated by normal distribution in MI. Categorical and continuous variables can be imputed using multivariate imputation by chained equations (Van Buuren et al. 1999), which is available in many statistical software packages. The statistical power of the analyses on the sub-samples of COURAGE or Physical Activity and Fitness sub-studies can be increased using the corresponding variables of the general Health 2011 Survey data and MI.

The participants of the COURAGE sample were asked not only the COURAGE-specific questions but also the Health 2011 Survey-specific questions. This means that there are variables, which measured the same phenomenon as a COURAGE variable and are therefore strongly associated, and variables, which measure similar phenomena and are associated with a COURAGE variable. As an example, consider questions EC05 (a to e) of the Health 2011 Survey: "*How many times have you during the past 12 months seen: a) a health centre dentist? b) a private dentist? c) some other dentist? (students' health care, defence forces, university, hospital etc.) d) dental technician? e) received some other dental treatment*?" cover the COURAGE question Q4068a (last 12 months) but only partially question Q4068b (last 2 weeks). In the latter case answers of zero times in EC05 would imply a value "*no*" in Q4068b, and MI would be applied in the cases of any positive values in EC05. Matching of the Health 2011 Survey and COURAGE variables has not been done in detail, thus we do not have yet any worked examples on the possible benefits, which the MI of the data on non-COURAGE participants could yield.

MI can, therefore, be applied to augment the missing COURAGE variables in the non-COURAGE sample. Especially in the case of the strong associations, the missing values can be

imputed with good predictive power. The inclusion of the non-COURAGE sample by means of the MI techniques can therefore improve the power of COURAGE analyses in many cases thus reducing the effects of non-response.

Clustering of the data

The original sample was based on a complex sampling design, which results in a geographically clustered data set. The clustering levels were based on the health centre districts. The clustering yielded considerable design effects especially in blood pressure measurements.

Repeated measurements on the same participants introduce a new level of clustering in the data. There is much more variance between participants than between, for example, the health centre districts. Therefore also design-based methods, in which only one clustering level can effectively be accounted for, can be applied in analysing the longitudinal Health 2011 Survey data, if individuals are used as the primary sampling units instead of the health centre districts. Perhaps the most plausible analytical approach are linear or non-linear mixed effects models, in which all levels of clustering can be handled using different random effects. The variances of these random effects can be compared to assess the levels, which are the most influential to the results.

Analyses of cross-sectional studies

An important objective in population studies is to determine the distribution of a variable in the population at a certain time point. Another objective is to assess if there have been changes in the population distributions over time. Often cross-sectional studies are based on independent samples of the population at different times, in which case the analyses are simple. In this case the same individuals have been measured twice, thus the comparisons of the population distributions are based on dependent samples. A further complication is that the individuals are 11 years older in 2011, which must be accounted for.

The population distributions have been compared using design-based analyses (Djerf et al. 2008) in which the individuals were considered as the PSUs. Additionally, the strata in year 2000 were used as stratification variables. In addition to unadjusted descriptive statistics, also

their model-adjusted versions based on predictive margins (Lee 1981, Graubard and Korn 1999, Djerf et al. 2008) were applied in order to adjust for confounders such as the different population distributions of age in years 2000 and 2011.

The cross-sectional results in 2000 and 2011 can be compared with other population studies containing other sample members than the Health 2000 and 2011 Surveys. For example, the Mini-Finland Survey can be analysed jointly, if the values of the variables describing the clustering levels (namely stratification, health centre districts and individual identification numbers) do not overlap in different surveys.

Analyses of longitudinal studies

In longitudinal studies individual changes and (baseline) factors related to them are of particular interest (Lynn 2009). Repeated measurements on same individuals can reveal changes much more efficiently than when using independent cross-sectional surveys. In estimating individual changes during the follow-up, it is likely that a researcher decides to retain baseline covariates, such as age, fixed.

The possibly informative right-censoring can complicate the estimation of individual changes over time. Individual changes cannot be estimated in a straightforward manner if only the survivors are considered when the variable of interest is associated with the risk of death. For example, the observed proportion of quitters was influenced by the smokers, who died between 2000 and 2011, thus the observed prevalence of quitters was likely to be an overestimate, because smokers generally have a higher risk of death than quitters. Crude descriptive analyses on categorical variables can be conducted considering death as an extra category of the variable of interest. For example, the smoking indicator with values smoker or non-smoker in year 2000 was extended to a variable with categories smoker, non-smoker or dead in year 2011. This kind of informative right-censoring should be accounted for especially if the study involves elderly people or other groups having a high risk to die. See, for example, an article by Stenholm et al. (2012) in which a selection model (Diggle and Kenward. 1994) and the Mini-Finland resurvey data of the Health 2000 Survey were utilized.

A group of the members of the Mini-Finland sample living in seven regions in 2000 were invited to the Health 2000 and 2011 Surveys. Therefore, we have a subset of individuals examined three times. Migration may have been associated with factors which were dependent on the outcomes of interest, thus there has been an informative selection process. There is a

need to analyse which factors predicted migration to or from the seven areas, and this information can provide us with more accurate results in both cross-sectional analyses and in individual follow-up studies.

HEALTH 2011 SURVEY – CONTENTS

In the following chapters, the contents of Health 2011 Survey are described in detail. The main focus is in the health examination which also included an extensive interview. The contents of interviews and questionnaires are illustrated in Appendices 4 and 5. All the forms used in the survey are also available on the website of the Health 2011 Survey (www.thl.fi/terveys2011).

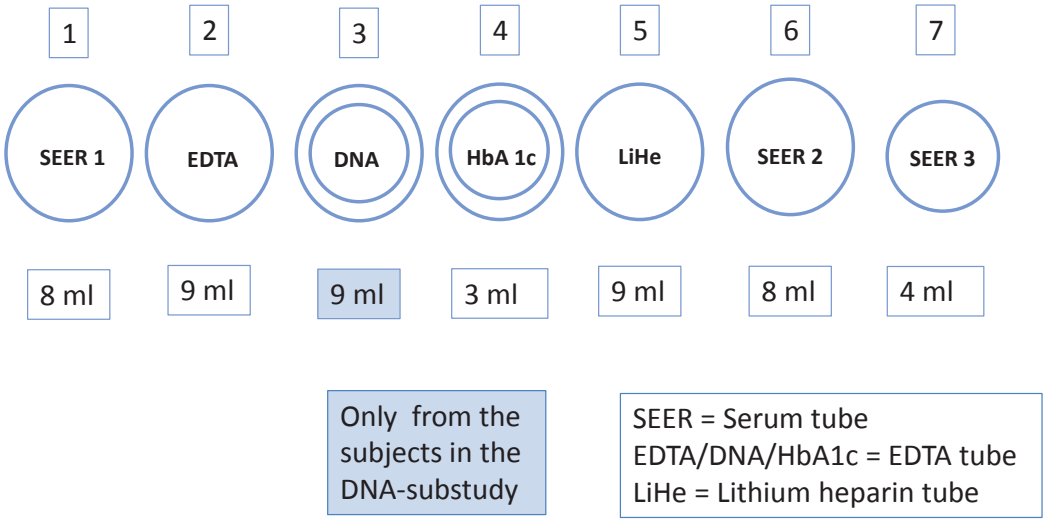
5. LABORATORY MEASUREMENTS

Laura Lund, Britt-Marie Loo, Jaana Leiviskä and Jouko Sundvall

5.1 Blood collection

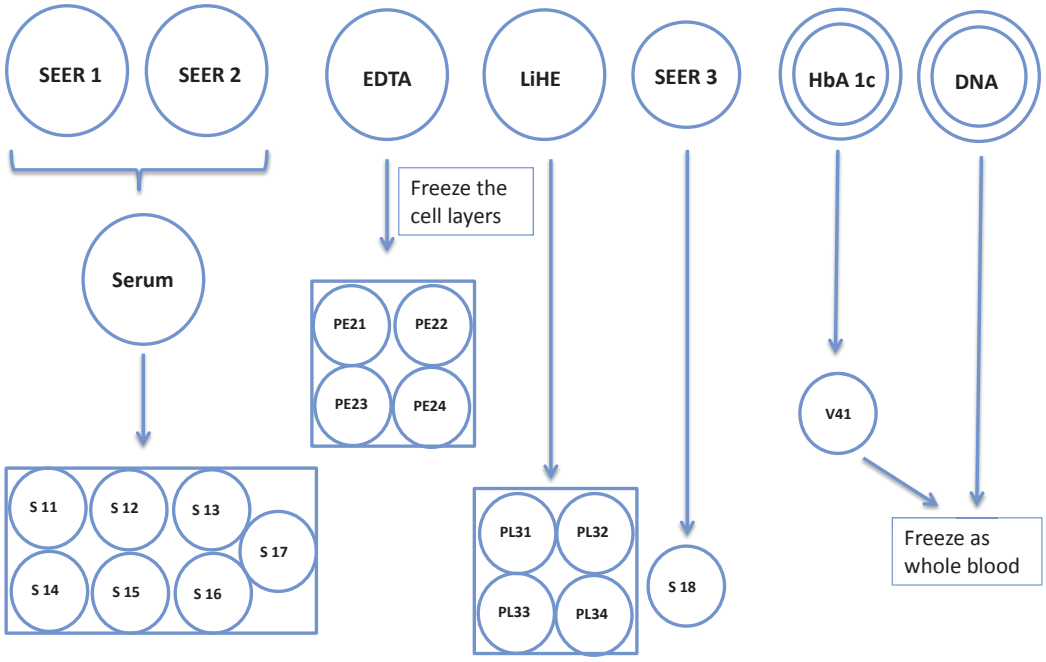
Samples of whole blood, serum, lithium heparin plasma and EDTA plasma were collected from all participants at health examination (Figure 5.1.1). The samples were divided into aliquot tubes as illustrated in Figure 5.1.2. The DNA sample was collected from participants belonging to the DNA sub-study (see Chapter 1.2). During the home-visit health examination, a serum sample and a DNA sample were collected only if there was access to a centrifuge. A flowchart of sampling and sample processing is shown in Figure 5.1.3.

Figure 5.1.1 Blood tube chart.



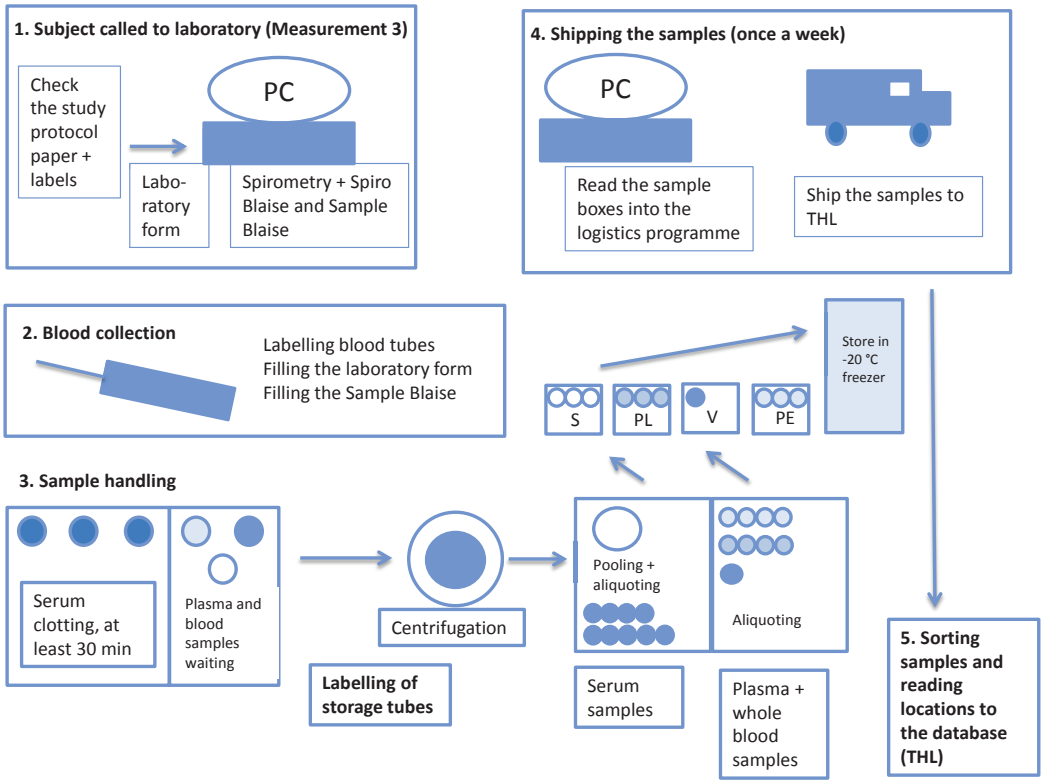
All the necessary supplies were delivered to the central field laboratories in advance. If there was lack of some equipment, it was ordered from THL and delivered to the sites. The field laboratory personnel estimated the amount of the supplies needed and delivered them to different study sites. All the required equipment, including a field centrifuge and a chest freezer, was moved with the laboratory. Electrical and manual pipettes were used in aliquoting the samples.

Figure 5.1.2. Aliquot tube chart.



Six to seven tubes of blood were drawn from each participant (5 to 10 ml plastic-walled, evacuated Venosafe tubes, gel tubes for serum). The sampling order was determined by the purpose of the samples. If difficulties were met with drawing the samples, three attempts were made with the participant's approval, to get at least one serum tube for basic laboratory measurements.

Figure 5.1.3. Flowchart of sampling and sample processing.



Venous blood samples were collected from a vein in the arm, with the participant in a sitting position. To prevent hemolysis the tourniquet was released as soon as the blood began to flow. If a sample could not be obtained from the arm, it was drawn from the back of the hand using a wing or open needle. Serum, EDTA and lithium heparin tubes were carefully inverted six times against the plugs.

After sampling, the nurse affixed labels to the tubes. The serum was allowed to clot for at least 30 minutes after the final tube had been collected. The maximum storage time allowed at room temperature before centrifugation was 60 minutes. The plasma samples were kept at room temperature for the same time as serum samples so that they could be centrifuged together.

An extra EDTA blood sample was collected from participants belonging to the DNA sub-study.

Participants were asked to fast at least for four hours before coming to the study. The last time the participant had eaten or had drunk anything but water was asked and recorded. If the fasting time was shorter than four hours, the length of the fasting and information on the consumed type of food or drink was recorded.

Information on plausible infections was also collected. The participants were asked if they had had any infections during the past seven days. The nurses were instructed to record all infections (respiratory or other type) that might have an influence on the CRP value.

5.2 Sample processing and management

The samples were processed according to Figures 5.1.2 and 5.1.3. One tube of EDTA whole blood from participants belonging to the DNA sub-study was frozen as such. The sample was frozen without opening the cap. From another EDTA blood tube (HbA1c tube) an aliquot of whole blood was pipetted and frozen. The serum, EDTA and heparin plasma tubes were centrifuged at 2200 G for 11 minutes. The sera from two centrifuged gel serum tubes were collected into one large pooling tube. The pooled serum was mixed by carefully inverting the tube five times and aliquoted into 1 ml tubes. If one of the serum tubes was haemolysed, it was not added to the pool but pipetted into separate aliquot tubes. The third serum tube (5 ml gel tube) was also centrifuged and the serum was separated into a 3 ml tube. Both EDTA and lithium heparin plasma tubes were centrifuged and pipetted into 1 ml aliquots. The white and red blood cells from the EDTA plasma tube were frozen in its primary tube. All the samples were frozen at –20 °C immediately after handling.

The date and time of sample handling, the ID code of the nurse, the number and type of samples obtained, and the volume of all serum and plasma aliquots were entered into the laboratory form. Any deviations in sampling or in sample processing were also recorded on the laboratory form. The forms were stored in electronic form using Blaise.

The participants handed the nurse their folder which included a study scheme form and labels for sample tubes, arranged in the order the samples were to be drawn. Each label in the sheet carried the same recurring secondary key for that particular set of labels. Storage tube labels also included an unequivocal primary key in both barcode and alphabetical format and a code

describing the type of sample. In the laboratory, the secondary key sticker was affixed to the study scheme form to mark that the participant had passed through the laboratory. The laboratory form was also labelled with a special label designed for this purpose. At the registration point, the label sheet was affixed to the health examination schedule of the participants to link the labels and the participant together.

5.3 Storage and shipment of samples

Serum, plasma and whole blood samples were immediately frozen at –20°C on site, normally within 45–60 min but not later than 120 min from sampling. If the samples were left at room temperature for longer than the maximum time (60 minutes) after the centrifugation, it was recorded to the laboratory form. The samples were stored in fibreboard boxes that had been labelled before they were put to the freezer. The storage boxes were filled with tubes according to a pre-planned box chart. The samples in the boxes were packed in dry ice and transferred from the field storage points to their final storage location at THL no later than 1–2 weeks after sampling. The sample boxes were packed in dry ice and shipped via a door to door carrier. The sample boxes were read to the logistic programme at the field laboratory before shipment to keep track of the boxes sent. A mechanical thermometer was also included in the transport container to monitor the temperature during transport.

When the sample shipment arrived at THL, the temperature of the shipment was checked and recorded. The content of the shipment was read into the logistic programme to make sure that all sample boxes that were sent from the field had also been received in THL. The serum, plasma and whole blood sample aliquots were sorted on the basis of the aliquot type and transferred to storage at –70 °C. DNA samples were stored at –20 °C.

5.4 Field laboratory quality assurance

Procedures for sampling and sample processing were tested during two pilot phases. Field laboratory personnel was trained in advance. The sampling and sample handling sites were audited once by an external auditor. Also the person responsible for the laboratory work audited the sites one or two times during the field examinations. Guidelines for field laboratory procedures were provided in *"Laboratory sampling process in the Health 2011 Survey"*

document. The instructions were completed with IT instructions, THL guidelines concerning the risk of infection, equipment manuals, operational safety bulletin for dry ice and contact information.

One of the major concerns during planning was how to minimize the risk of errors during sampling and sample handling. Therefore all the samples from each participant were processed at the same time. In the event of problems, the field laboratory personnel contacted the person in charge at THL by e-mail or phone. E-mail communications during the field stage and reports on errors and problems were filed and archived. A job rotation scheme was operated with personnel responsible for sample collection, sample processing and the duties of interview nurse, exchanging jobs at about one-week intervals.

5.5 Basic laboratory measurements

The basic laboratory measurements of alanine aminotransferase, apolipoproteins A-I and B, calcium, cholesterol, creatinine, glutamyltransferase, glucose, HDL-cholesterol, high sensitive CRP, triglycerides and uric acid were performed at the Disease Risk Unit at THL, Helsinki. LDL cholesterol was calculated using the Friedewald formula. The measurements were performed on a clinical chemistry analyser Architect ci8200 (Abbott Laboratories, Abbott Park, IL, USA). The testing laboratory of the Disease Risk Unit (No. T077) is accredited by the Finnish Accreditation Service, FINAS and it fulfils the requirements of the standard SFS-EN ISO/IEC 17025:2005. The scope of accreditation covers all analyses except uric acid. The determinations were carried out on frozen serum samples within two months of sampling. Table 5.5.1 provides more detailed information concerning the methods used. The data produced in later analyses will be added into a catalogue of laboratory methods.

For standardising the measurements, the laboratory has taken part in the Lipid Standardization Program organised by CDC, Atlanta, USA and External Quality Assessment Schemes organised by Labquality, Helsinki, Finland. The quality of the results of the series of analysis was ascertained by using controls, which were used to determine interassay coefficients of variation (CVs). During the course of the study comprising six months in 2011 and 2012, the precision between series expressed as coefficients of variation (CV%) and the accuracy of the methods (mean bias% ± SD) are demonstrated in the Table 5.5.1. The bias indicates the difference between the laboratory's own result and the target value of the quality assessment sample and describes the laboratory's systematic error.

Table 5.5.1. The accuracy of the methods, differences between the laboratory's own result and the target value.

Assay	Method	CV% ± SD	Bias% ± SD
Cholesterol	Enzymatic CHOD PAP, Abbott	0,6% ± 0,1, CDC 4/2011 0,8% ± 0,1, CDC 1/2012 TLAB-contr, N=25, mean=3,84 mmol/l, CV%=0,9	CDC's Lipid Standardization Program (Atlanta, USA) 4/11 – 1/12 0,2% ± 0,3 (N=6)
HDL-cholesterol	Homogenous, Accelerator Selective Detergent, Abbott	1,2% ± 0,1, CDC 4/2011 1,7% ± 0,8, CDC 1/2012 TLAB-contr, N=25, mean=1,21 mmol/l, CV%=1,7	CDC's Lipid Standardization Program (Atlanta, USA) 4/11 – 1/12 1,6% ± 1,8 (N=6)
Triglycerides	Enzymatic GPO PAP, Abbott	1,3% ± 0,3, CDC 4/2011 1,5% ± 0,4, CDC 1/2012 TLAB-contr, N=25, mean=0,83 mmol/l, CV%=2,1	CDC's Lipid Standardization Program (Atlanta, USA) 4/11 – 1/12 -5,8% ± 1,8 (N=6)
Alanine Aminotransferase	NADH (with P-5'-P), IFCC, Abbott	BIO45610L1, N=167, mean=16,7 U/l, CV%=5,5 BIO45610L3, N=167, mean=195 U/l, CV%=2,5 TLAB-contr, N=25, mean=16,9 U/l, CV%=7,6	Labquality's General Clinical Chemistry Program (Helsinki, Finland) 5/11 - 1/12 0,8% ± 5,5 (N=6)
Apolipoprotein A1	Immuno-turbidimetric, Abbott	1,3% ± 0,3, CDC 4/2011 1,5% ± 0,0, CDC 1/2012 TLAB-contr, N=25, mean=1,38 g/l, CV%=1,2	CDC's Lipid Standardization Program (Atlanta, USA) 4/11 – 1/12 -2,9% ± 3,3 (N=6)
Apolipoprotein B	Immuno-turbidimetric, Abbott	1,5% ± 0,3, CDC 4/2011 2,0% ± 0,9, CDC 1/2012 TLAB-contr, N=25, mean=0,65 g/l, CV%=2,2	CDC's Lipid Standardization Program (Atlanta, USA) 4/11 – 1/12 -9,5% ± 2,3 (N=6)
Calcium	Arsenazo III Dye, Abbott	BIO45610L1, N=214, mean=1,52 mmol/l, CV%=1,6 BIO45610L3, N=213, mean=3,29 mmol/l, CV%=1,0 TLAB-contr, N=26, mean=2,38 mmol/l, CV%=1,1	Labquality's General Clinical Chemistry Program (Helsinki, Finland) 5/11 - 1/12 -2,1% ± 1,4 (N=6)

Table 5.5.1. Continued.

Assay	Method	CV% ± SD	Bias% ± SD
Creatinine	Enzymatic, Abbott	BIO45610L1, N=177, mean=57 µmol/l, CV%=2,8 BIO45610L3, N=175, mean=581 µmol/l, CV%=1,2 TLAB-contr, N=25, mean=82 µmol/l, CV%=2,5	Labquality's General Clinical Chemistry Program (Helsinki, Finland) 5/11 - 1/12 -1,9% ± 4,7 (N=6)
High Sensitive C-Reactive Protein	Immuno-turbidimetric, Sentinel	BIO45610L1, N=167, mean=0,37 mg/l, CV%=9,4 LIP57220L1, N=158, mean=0,87 mg/l, CV%=5,9 LIP57220L2, N=168, mean=5,2 mg/l, CV%=3,8 TLAB-contr, N=25, mean=3,5 mg/l, CV%=5,1	LabQuality's, Myocardial Markers Program (Helsinki, Finland) 1/11 - 2/12 5,9% ± 5,6 (N=4)
Gamma-Glutamyl Transferase	L-Gamma-glutamyl-3-carboxy-4-nitroanilide Substrate, Abbott	BIO45610L1, N=172, mean=29 U/l, CV%=2,9 BIO45610L3, N=171, mean=135 U/l, CV%=2,9 TLAB-contr, N=25, mean=23 U/l, CV%=3,3	Labquality's General Clinical Chemistry Program (Helsinki, Finland) 5/11 - 1/12 -5,9% ± 3,6 (N=6)
Glucose	Enzymatic Hexokinase, Abbott	BIO45610L1, N=201, mean=3,2 mmo/l, CV%=1,9 BIO45610L3, N=199, mean=20,4 mmo/l, CV%=1,7 TLAB-contr, N=25, mean=4,8 mmo/l, CV%=1,3	Labquality's General Clinical Chemistry Program (Helsinki, Finland) 5/11 - 1/12 2,8% ± 1,1 (N=6)
Uric Acid	Enzymatic Uricase, Abbott	BIO45610L1, N=175, mean=191 µmol/l, CV%=3,5 BIO45610L3, N=173, mean=554 µmol/l, CV%=2,2 TLAB-contr, N=23, mean=361 µmol/l, CV%=1,9	Labquality's General Clinical Chemistry Program (Helsinki, Finland) 5/11 - 1/12 -0,1% ± 3,8 (N=6)

CV = interassay coefficient of variation; SD = standard deviation; TLAB = THL, laboratory

To assess the analytical differences between laboratory methods in Health 2000 and Health 2011, the samples from about 240 participants participating in both surveys were chosen for

repeated analyses. The samples were analysed at the Disease Risk Unit and the Population Studies Unit. Differences between total, HDL and LDL cholesterols, triglycerides, apolipoproteins A-I and B, glucose, glutamyltransferase, creatinine, C-reactive protein, calcium and uric acid levels were estimated by regression analysis performed with the standardized principal component analysis (Table 5.5.2). The Pearson correlation coefficient describes the correlation between measurements of samples from the same individuals collected in the Health 2000 Survey analysed in 2001 and again in 2012. The regression equations can be used to transform results from the Health 2000 to the same level as the results in the Health 2011.

Table 5.5.2. The results obtained from samples collected in Health 2000 compared to results from samples of the same individuals collected in Health 2011.

Analysis	Regression equation	r	N
Cholesterol	Health2011 = 0,93016 × Health2000 + 0,07961	0,9954	238
HDL-Cholesterol	Health2011 = 0,95901 × Health2000 + 0,22417	0,9580	237
Triglycerides[1)]	Health2011 = 0,94362 × Health2000 – 0,01934	0,9984	236
Glucose	Health2011 = 1,01796 × Health2000 – 0,18989	0,9930	237
LDL-Cholesterol[2)]	Health2011 = 0,87709 × Health2000 + 0,06843	0,9921	226
Apolipoprotein A1	Health2011 = 0,93331 × Health2000 – 0,01468	0,9724	238
Apolipoprotein B	Health2011 = 0,85631 × Health2000 – 0,06485	0,9694	238
Glutamyltransferase	Health2011 = 1,11370 × Health2000 – 1,76395	0,9986	238
C-reactive protein[3)]	Health2011 = 1,01123 × Health2000 + 0,97208	0,9948	70
Creatinine	Health2011 = 0,99290 × Health2000 – 1,11654	0,9879	237
Uric acid	Health2011 = 1,06631 × Health2000 – 29,07657	0,9715	238
Calcium	Health2011 = 1,06312 × Health2000 – 0,13074	0,8408	237

r= Pearson correlation; N = number of observations

[1)] A single sample with exceptionally high triglyceride content (>11 mmol/l) was excluded from the analyses.

[2)] LDL-cholesterol was calculated by the Friedewald formula (LDLchol = Chol – HDLchol – 0,45 × Trigly). All samples with triglyceride concentration > 4 mmol/l were excluded from the calculations.

[3)] The regression analysis was performed on samples with CRP concentrations ≥ 1,2 mg/l. Due to differences in specificity and sensitivity between the methods used in 2003 and 2012, concentration results < 1,2 mg/l are not possible to convert.

6. ANTHROPOMETRIC MEASUREMENTS

Noora Ristiluoma, Johanna Mäki-Opas, Ulla Laitinen and Markku Heliövaara

All anthropometric measurements were carried out according to European standard protocols (Tolonen 2013). Some adaptation was necessary for home visits.

6.1 Height and weight

Height was measured by using a portable, stand-alone stadiometer (Seca 213). On assembly and every morning a carpenter's level was used to verify the correct vertical and horizontal placement of the stadiometer. When possible, the stadiometer was taped on the floor and on the wall. Alternatively the right placement of the stadiometer was marked with tape.

Height was measured with light socks or barefoot. Hair ornaments interfering with the measurement were removed. Participants were instructed to stand upright with the feet together on the stadiometer´s platform with the back of the head, back, buttocks and heels against the measuring rod. The head was positioned so that the top of the external auditory meatus (ear canal) and the bony orbit (cheek bone) were in a straight line. The measurement was read at eye level and recorded to an accuracy of 0.1 cm. Height was not measured if the participant was unable to stand upright or if the participant exceeded the maximum height of the stadiometer (205 cm). In these cases self-reported height was asked and recorded.

Weight was measured as a part of the bioimpedance body composition analysis (Seca 514, see Chapter 6.3). If the bioimpedance measurement was not possible, weight was measured using a digital floor scale (Seca 877). A small mechanical bathroom scale (Salter 409) was also used during the home visits. Weight was measured in light clothing and without shoes. The result was recorded to an accuracy of 0.5 kg.

Weight was not measured in very rare cases if the participant was unable to move, had difficulty in standing steadily or if the participant's weight exceed the maximum of the scale

(Salter 130 kg, Seca 200 kg). In these cases the participant's self-reported weight was recorded. The scales were calibrated every two weeks by comparing results with the bioimpedance analyser.

After the nurse had entered the height and weight values, BMI was calculated by the recording program. The participants received written and, if needed, oral feedback on their BMI.

Information on self-reported height and weight was collected by the self-administered Questionnaire 1 which the participants filled in before the health examination (see Appendix 5). Questions regarding self-reported height and weight were also included in the phone interview, short questionnaire and questionnaire for young adults (see Appendices 4 and 5). These allow a comparison between self-reported and measured values and assessment of BMI also for those who did not participate in the health examination.

6.2 Waist circumference

Before the measurement the participant was asked to expose the waist by loosening the belt, lowering the pants/skirt and lifting the shirt. The measurement was done on the bare skin using a flexible, non-elastic push-button measuring tape. If the participant felt uncomfortable or refused to undress, the measurement was taken over light clothing.

The participants were asked to stand with their weight evenly balanced on both legs, a small gap between the legs and hands hanging loosely beside the body. The person taking the measurement was seated in front of the participant. Waist circumference was measured at the midway between the lower rib margin and the iliac crest with the tape all around the body in horizontal position. The measuring tape was fastened firmly but so that the measurer´s finger was able to fit between the subject's body and the tape. The participant was instructed to breathe normally and the reading was taken during light expiration. The reading was recorded to the nearest millimetre.

For quality control, the length of the flexible measuring tape was weekly measured against a stiff metallic tape measure. Stretched measuring tapes were replaced. In addition, all measuring tapes were once a month replaced by new ones.

6.3 Body composition

Bioelectrical impedance analysis (BIA) is a rapid and non-invasive method for evaluating body composition. BIA is based on the fact that lean muscle tissue contains high levels of water and electrolytes, and therefore acts as a conductor of an electrical signal. Fat mass is comparatively anhydrous and acts as a resistor to the flow of an electrical signal. Increasing levels of fat mass result in a higher impedance value, and correspond to higher levels of body fat.

The Health 2011 Survey opted to use a new device (Seca mBCA Model 5154 with Software Seca analytics 115, Seca, Hamburg, Germany), which weighs the participant, measures resistances and impedances at frequencies of 1, 1.5, 2, 3, 5, 7.5, 10, 15, 20, 30 50, 75, 100, 150, 200, 300, 500, 750 and 1000 kHz, saves all the data and – working on the assumption that the body consists of five cylindrical segments (four limbs and the body) – computes fat and fat-free masses immediately (Bosy-Westphal et al. 2013). Further estimates of body composition, as total body water (TBW), extracellular water (ECW) and lean soft tissue (LST) in five segments were computed after the fieldwork.

Each study participant was prepared for the measurement by wiping the soles of the feet and the palms. This served the dual purpose of cleaning and disinfecting as well as improving electrical conductivity. The participant was instructed to step on the analyser and to touch the electrodes. The device includes a hand-rail with three electrode pairs. Depending on the height and arm length of the participant the appropriate electrode pair was chosen. The participant's position was adjusted if necessary. The nurse entered the participant's data, which showed up on the analyser display. Age was entered to an accuracy of one year and body height to an accuracy of one centimetre. Sex was indicated with a separate key. After the 'Start' button was pressed, the measurement lasted about one minute. The nurse offered a brief interpretation of the findings. The report automatically produced by the device was printed and a copy was given to the participant. The measurement was not carried out if the participant had a cardiac pacemaker.

Not all BIA measurements were successful. Especially during the first few weeks of the fieldwork, there were some technical problems. Overall the number of participants not tested was, however, small.

In the Health 2000 Survey another eight-polar BIA device (InBody 3.0, Biospace, Soul, South Korea) which measured segmental resistances and impedances only at 5, 50, 250 and 500 kHz, was used. The agreement is satisfactory but small systematic shifts between the devices cannot

be ruled out. In general, the BIA methods have been validated (Bedogni et al. 2002, Thomas et al. 2003, Jaffrin 2009, Ward 2012). The results of BIA from the Health 2000 and Health 2011 Surveys have great potential application. This concerns particularly the segmental estimates of fat-free mass, which cannot practically be estimated using other anthropometric measurements.

7. CLINICAL MEASUREMENTS

7.1 Blood pressure and heart rate

Antti Jula and Ulla Laitinen

Blood pressure was measured in the health examination after the anthropometric measurements. Participants were requested to refrain from heavy exercise and to avoid cola drinks, coffee, tea, eating, and smoking for at least one hour before the measurement.

Blood pressure was measured from the right arm whenever possible. On arrival to the blood pressure measurement, the participants were seated and asked to undress their upper body so that the upper right arm was bare for the blood pressure cuff and no clothes restricted blood flow in the arm. If the right arm was amputated or if proper readings could not be obtained for other technical reasons, the measurements were taken from the left arm.

Current instructions (Rose et al. 1982, Finnish Hypertension Society Working Group 2002) and the European standard protocols (Tolonen 2013) were followed in checking the sitting position of the participant, resting before and between the measurements, wrapping the cuff around the upper arm, positioning the cuff at the level of the heart and in listening to the Korotkoff sounds.

Measurements were performed with a standard mercury manometer (Riester Diplomat Presameter Desk Set). All manometers were serviced and calibrated before use. The proximal circumference of the upper arm was measured at a height of 5 cm above the elbow crook of the arm. If the circumference was 35 cm or less, a standard cuff (bladder width 12 cm, length 35 cm) was used. For a circumference exceeding 35 cm, a larger cuff (bladder width 15 cm, length 43 cm) was used.

After fastening the cuff, the participants were sitting calm for five minutes before beginning the measurements. Measurements began by determining the level of the systolic pressure. This was conducted by steadily pumping air into the cuff while palpating the wrist artery until the pulse

could no longer be felt. The point at which the pulse disappeared was noted by the observer and the pressure from the cuff was released.

For the first blood pressure measurement, the bell of the stethoscope was placed at the crook of the arm at the point where the pulse was strongest. By pumping air into the cuff, the pressure was raised to 30 mmHg above the systolic blood pressure level determined earlier. The pressure was then released steadily at a rate of approximately 2 mmHg per second. Systolic pressure was recorded at the appearance of the first Korotkoff sounds to an accuracy of 2 mmHg. Diastolic pressure was also recorded to an accuracy of 2 mmHg at the fifth phase of the Korotkoff sounds, when the latter of two consecutive sounds was no longer audible.

After the first blood pressure measurement pulse was measured by palpating the wrist artery and counting the number of pulses from the artery for 60 seconds. The second blood pressure measurement was then conducted following the same method as in the first measurement. After both measurements the nurse informed the participant about the results and gave instructions to seek professional help if needed. The blood pressure values were also recorded on the feedback form for all participants.

To ensure favourable measurement conditions, measurement room temperature was recorded twice each day: in the morning before the first subject arrived and in the afternoon after the last participant had left.

The quality of blood pressure measurements was ensured by providing both theoretical and practical training on blood pressure measurement for the nurses. The nurses were trained in pairs and at the end of the training all nurses took part in a skill test. During the fieldwork, the quality of blood pressure and heart rate measurements was constantly monitored by following individual score variation as well as the variation of values attained by the two nurses of the same field team and the variation between the teams. Key experts visited also all field teams to ensure that the correct protocol was followed. Detailed information on quality of blood pressure measurements is available in Chapter 4.2.

Blood pressure was measured at home during the nurse visits and for the Finn-Home study by the participants themselves with a validated automatic oscillometric device (Omron model M& HEM-7211-E(V), Omron Corp.-, Kyoto, Japan) according to the protocol of the Finn-Home study (Reunanen 2008).

7.2 Resting ECG

Hannu Karanko

Routine 12-lead resting ECG was recorded with GE MAC 5000 or MAC5500 (GE, Freiburg, Germany), computerized recording system, for each participant. All five recorders had equal settings. Electrical impulses were recorded (5+5 sec) and stored in digital memory and two printouts on paper with six lead views each were taken. The hard-copy was provided with ID information (see below), a short measurement matrix measured from summary complexes and an automatic diagnostic assessment (12SL-system, Marquette Electronics Inc., Milwaukee, WI, USA). All recorders were programmed to store and print the ECGs pressing once the recording button. A recorder specific location number was programmed to separate different measurers and field work teams as well.

After some structural measurements, the participants were asked to lie down comfortably on an examination bed, cloths removed from the chest and ankles. Skin preparation included hair removal from chest, if necessary, and sometimes light scraping with special sand-paper (3M Red Dot, Trace Prep 2236, 3M Canada). No chemicals were used. After preparation of skin, 10 disposable electrodes (GE, Silver Mactrode Plus) were placed: four on extremities and six on the chest. If the participant had an amputated limb, the electrode was placed as far as possible on the stub. Electrodes were positioned in accordance with instructions for standard 12-lead ECG measurements (Heikkilä 1982, Rose et al. 1982). Ten lead outlet boxes were connected to the corresponding electrodes. Surname, ID code, age and sex as well as height and weight, measured just before the ECG recording, were fed into the recording system. No information on medication or blood pressure was entered.

After the above preparation, participants were asked to relax, not to move or speak and to breathe lightly and regularly. The quality level of recording could be seen on the instrument's screen and when no more baseline wandering or muscular or other artefacts were present, the recording button was pressed once. Normally, no built in disturbance elimination was used. After this, the hard-copy was checked and, if satisfactory, a second paper copy was taken. If not satisfactory, the disturbance was eliminated and the recording was repeated. If the second recording was technically satisfactory, the first one was deleted and paper copies were disposed of. One copy was attached to the participant's paper file and the other was given to the participant. Paper speed was 50 mm/s and the filter was set at 150 Hz.

Fieldwork nurses were trained to take and to read ECG recordings and diagnostic suggestions. They were also provided with written information on 12SL diagnostic suggestions to be able to assess the severity and urgency of a possible abnormality. The assessment led to a three grade classification, which correlated to the significance and urgency of the abnormality. The nurses sent the most urgent cases to the local health centre for follow-up and treatment within the same day. In less urgent cases they had an opportunity to consult the survey physician.

All five recorders stored ECGs on diskettes or memory cards. Information stored on removable media was copied to the hard disc of a laptop computer once a day. New memory media was inserted after about one hundred recordings were stored. Diskettes and memory cards provided with directory list were sent to THL office in Turku. Laptop files served as security copies of the ECG recordings.

In THL office in Turku the contents of diskettes were stored to the local MUSE network server system (MUSE CV, Marquette Electronics Inc., Milwaukee, WI, USA). The storing procedure included, if necessary, corrections of ID code, surname, sex and age. Some location corrections were also done. Directory lists were used to add handwritten information if the participant's diagnostic information was wrong or misleading or if the measurement points for different cardiac intervals were not correctly placed. The checked and corrected ECG data were analysed with the Magellan software program (Marquette Electronics Inc, Milwaukee, Wi. USA) to create Excel files with durations and amplitudes of P waves, T waves, QRS complexes, durations of conduction times (QT and PR intervals) and morphology of the ST-segment of every single 12-lead ECG.

7.3 Spirometry

Ulla Laitinen, Markku Heliövaara and Tuula Vasankari

Spirometry measures ventilation, the movement of air into and out of the lungs, the nature of any lung dysfunction (bronchial obstruction and restriction of the lungs), and its severity. Spirometry is primarily used for diagnosing respiratory diseases, monitoring the effectiveness of treatment and assessing the severity of disease (Cotes 1975). Spirometry tests are therefore essential in the assessment of asthma and chronic obstructive pulmonary disease. Flow-volume spirometry is a sensitive method that may allow for early detection of changes in the respiratory tract caused by smoking. In population studies spirometry has been used for

monitoring respiratory inadequacies and measuring respiratory function (Aromaa et al. 1985, Heliövaara 2008)

In the Health 2011 Survey, spirometry measurements were taken by Medikro´s® Spirometry System which included the Medikro® SpiroStar flow-volume spirometer and Medikro ® Spiro2000 software. Disposable SpiroSafe flow transducers were also used. The spirometer was calibrated with a one-litre calibration pump, and the equipment was checked every morning before the measurements.

The key measurements taken in the Health 2011 Survey included:

- forced vital capacity (FVC)
- forced expiratory volume in one second (FEV1)
- forced expiratory volume in one second, percentage (FEV1/FVC, FEV%)

If obstruction (FEV1/FVC < 70%) was detected, a bronchodilator test was performed to determine reversibility.

Before each measurement, the nurse filled in the participant's personal information including last name, the personal ID code, date of birth, sex, height and weight into the spirometry software. The participant was instructed to sit on a chair holding a good posture. A disposable mouthpiece was placed into the participant's mouth between the teeth, lips firmly around the mouthpiece and a nose clip was used to ensure the flow of respiration through the mouth. As the measurement began, the nurse instructed the participant to breathe in and out normally a couple of times and then to fill their lungs with air to maximum capacity. When the lungs were full, the participant was instructed and urged to exhale as forcefully as possible and for as long as possible until the lungs felt empty of air.

At least two measurements were carried out by each participant as the aim was to produce two as consistent curves as possible. The maximum permissible difference between the two highest FEV1 and FVC values was 10 per cent. This often required several efforts. The participant was allowed to rest and to remove the mouthpiece between the measurements. After two consistent curves were obtained, the software automatically selected the best measurement and the corresponding FEV1, FVC and FEV1/FVC values were recorded into the file of the participant. The largest FEV1 and FVC were selected from these two consistent curves, even if they did not come from the same curve.

A bronchodilator test was performed if the participant showed good motivation and if the test performance as such was satisfactory but FEV1/FVC was less than 70%. For the test, the nurse administered two 0.1 mg doses of salbutamol aerosol (Ventoline Evohaler 0.1 mg/dose) into a holding chamber (Volumatic) and made sure that the mouthpiece was firmly and tightly in the participant's mouth. The participants were instructed, after a normal rest exhalation, to steadily fill their lungs with air through the holding chamber and then to hold their breath for five seconds. The spirometry test was then repeated a minimum of 15 minutes after the administration of salbutamol. The FEV1, FVC and FEV1/FVC values were recorded as before.

After the measurements, the participants were given a printed copy of their measurement results and oral information on their test results. Co-operation of the participant was recorded only if it was poor or otherwise unusual. If no measurements were obtained, the nurse recorded the reason into the file of the participant.

Quality control and comparability of the results

The success of spirometry and the reliability of its results depend most crucially on full adherence to the procedural guidelines set out. The results for FEV1/FVC in the Health 2011 Survey are highly comparable with the results from the Mini-Finland Survey and the Health 2000 Survey, which indeed was the main objective in designing the measurements. In the pilot study, no substantial difference was observed between the devices used in the Health 2000 Survey and those used in the Health 2011 Survey (Table 7.3.1). The Mini-Finland Survey did not, however, include bronchodilator tests.

Test-retest reliability coefficients for FVC and FEV1 values for the spirometers used in these two surveys were within the range of 0.93–0.97. However, spirometric measurements of breathing function depend not only on the health of the respiratory system (Cotes 1975), but also on the mechanical properties of the chest, muscle function and the condition of teeth or dental prostheses. All these must be taken into account in interpreting the results.

Table 7.3.1. Comparison between two spirometry devices, SpiroStar (M947900100635) and Vitalograph (Type 2150, serial number VM000410), at the pilot of Health 2011 Survey, May 2011, among 38 volunteers.

	SpiroStar		Vitalograph			Comparison	
	mean	SD	mean	SD	α	mean difference	p
FEV1	3.29	0.73	3.32	0.71	0.93	-0.04	0.37
FVC	4.2	0.93	4.34	0.91	0.97	-0.07	0.03
FEV1/FVC %	77.47	7.68	77.06	8.18	0.77	0.41	0.60

SD = standard deviation; α = reliability coefficient; p = p value for systematic shift; FEV1 = forced expiratory volume in one second; FVC = forced vital capacity; FEV1/FVC % = forced expiratory volume in one second, percentage

8. SOCIODEMOGRAPHIC FACTORS AND LIVING CONDITIONS

Seppo Koskinen and Tuija Martelin

8.1 Sociodemographic factors

Information on age, sex, date and place of birth, marital status, place and type of residence and mother tongue was obtained for the whole sample (participants and non-participants) from the Population Register Centre. In addition register data were acquired from Statistics Finland concerning level of education, occupation and socioeconomic position of on all individuals included in the sample. Finally, the Ministry of Employment and Economy provided information on unemployment (see Chapter 4.3).

At the beginning of the interview the participants were asked about their current and previous living arrangements, education, main activity, employment and unemployment history, characteristics of the current/latest job and spouse's main activity (see Appendix 4). Some key questions were also included in the phone interview (see Appendix 4) as well as in the short questionnaire and the questionnaire for young adults (see Appendices 4 and 5).

8.2 Financial situation

Information on the financial situation of the respondents was gathered by means of posing a question "*How would you describe the current balance between income and expenditure in your household?"* which was included in Questionnaire 1 (see Appendix 5). The respondents were asked to evaluate their situation by choosing one of the following responses: *1) We have more than enough money to cover our needs; 2) There is enough money to cover our needs; 3) We have to some extent to compromise when deciding what we do with the money; 4) We have to compromise considerably in our consumption but we can manage with our income; 5) We*

have to make major compromises in our consumption and despite of that we do not manage with our own income; 6) I cannot say/it is hard to estimate. An identical question was used in the Health 2000 Survey.

8.3 Living conditions in childhood

Information on living conditions in childhood was collected among young adults. In the interview for young adults, retrospective information was collected on the parental educational level and parental occupation as well as on living arrangements in childhood (see Appendix 4). Questionnaire 1 for young adults included questions on childhood adversities, long-term financial problems in the family, frequent parental unemployment, parental divorce, parental alcohol problems, parental mental health problems and parental serious illness or disability, respondent's own serious chronic disease, and having been bullied at school (see Appendix 5). Also those young adults who only received a postal questionnaire for young adults were asked most of the above questions (see Appendix 5). In the Health 2000 Survey, subjects aged 29 years or older had been asked corresponding retrospective questions.

8.4 Living environment and housing

Information on living environment and housing was obtained both through Questionnaire 1 and the interview. Questions concerning safety were the same that have been used in earlier Finnish studies on safety and insecurity (Heiskanen et al. 2000). In the questionnaire, participants were first asked whether they feel unsafe when walking in their neighbourhood and secondly whether they are afraid to be alone outdoors in the evenings after 10 PM. Moreover, one of the items of the EUROHIS-QOL 8-item index (see Chapter 17) concerned living environments.

In the interview, there were questions on place of residence comprising questions on type of residence (private residence, sheltered housing unit etc.) and – for respondents aged 55 years or over – on distance to and way of commuting to the nearest food store, pharmacy, health centre and to neighbours, friends or relatives. In addition, questions on disadvantages, facilitating factors and safety equipment in housing were asked from subjects aged 70 years or over. The COURAGE sub-study included a large section on the assessment of the built environment (see Chapter 19.2).

9. WORKING CONDITIONS, WORK LOAD AND WORK ABILITY

The items concerning work and work ability were based on questions in several earlier surveys (Aromaa et al. 1989, Piirainen et al. 2000, Tuomi et al. 2006, Gould et al. 2008, Koskinen et al. 2012). Information was collected by questionnaires and interview (see Appendices 4 and 5).

9.1 Working conditions

Päivi Sainio and Seppo Koskinen

Questions on perceived physical and chemical hazards at work were asked in the interview from all respondents who had worked at any time during the past 12 months (see Appendix 4). The set of questions concerned noise, dust, vibration, chemicals, gases, cigarette smoke, coldness, heat, air draught and insufficient lighting at work. The response options ranged from 1=not present or no harm to 4=hinders very much.

9.2 Physical work loads

Eira Viikari-Juntura and Svetlana Solovieva

History of physical work exposures associated with current and previous jobs was collected during the interview (Viikari-Juntura et al. 1996, Solovieva et al. 2012). The participants were asked to list their current and previous occupations between 2000 and 2011, in which they had worked for at least one year. Information on physical workloads was collected for the most recent occupation and a maximum of five previous occupations in which the respondent had worked for the longest periods. First, the physical loading in the current and previous jobs was asked with the following general question: *"Is the job in question physically demanding, involving e.g. lifting and carrying heavy loads, excavating, shovelling or hammering?"* After

this, the following dichotomous questions were put: *"Does the job in question involve kneeling or squatting for at least one hour per day?", "driving a car, tractor or other motor vehicle for at least four hours per day?", "manual lifting, carrying or pushing items heavier than five kg at least twice per minute for at least two hours per day?", "manual lifting, carrying or pushing items heavier than 20 kg at least ten times per day?", "working with hands above the shoulder level for at least one hour per day?", "working in a forward bent position without support for at least one hour per day?", "work demanding high hand grip forces (for example, squeezing, twisting, holding burdens or tools) corresponding with a grip force of 3 kg for at least one hour per day?", "repetitive movements of the hands or wrists (for example packing and sorting out) for at least two hours per day?", "keying (for example typing, cashier work, visual display work) for at least four hours per day?", "working with a vibrating tool for at least two hours per day?", "work that requires sitting (excluding machine operating and car driving) for at least five hours per day?" and "work that requires standing or walking at least five hours per day?".* In addition, the respondents were asked about the timing (starting and ending year) of the jobs in question.

One question about physical demands at work was included also in Questionnaire 1 (see Chapter 10.4).

9.3 Psychosocial working conditions

Kirsi Ahola

Data on psychosocial working conditions (Karasek et al. 1998) and burnout (Maslach et al. 1996) were collected in Questionnaire 1 (see Appendix 5).

Psychosocial working conditions included job strain, social support, and job insecurity. Job strain was assessed using items regarding job demands and job control from the Job Content Questionnaire (JCQ, Karasek et al. 1998). To create an indicator of job strain, the job demand and job control scales can be dichotomized at their median and the following sub-groups formulated: low strain (low demands and high control), active work (high demands and high control), passive work (low demands and low control), and high strain (high demands and low control).

Social support was assessed using two items from the Job Content Questionnaire (JCQ, Karasek et al. 1998): support from supervisor and support from colleagues.

Job insecurity was assessed by five questions estimating the threat of long-term unemployment, dismissal, decrease of tasks, or transfer to another job (Lehto 1991).

Burnout was assessed using the Maslach Burnout Inventory – General Survey (MBI-GS, Maslach et al. 1996, Kalimo et al. 2006) which is described in Chapter 12.

9.4 Work ability

Raija Gould, Päivi Sainio and Seppo Koskinen

The questions concerning work ability were presented in the interview (see Appendix 4: IB-section), and they were identical to those included in the Health 2000 Survey (Aromaa and Koskinen 2004). The instruments are described in more detail by Gould et al. (2008) in the report *"Dimensions of Work Ability"*, and in the TOIMIA-database (in Finnish, Gould et al. 2015).

Work ability estimate

In the three-level assessment of work ability, the participants under age 75 years were asked to assess their current work ability regardless of whether they worked or not. If the respondents had trouble answering, they were asked to assess their work ability in relation to their most recent job. The options were 1=completely fit for work, 2=partially disabled for work, and 3=completely disabled for work. The question originates from the Mini-Finland Survey (Aromaa et al. 1989). Those with partial disability were asked to describe the restrictions in their work ability. Those completely unable to work were asked when they had become unable to work and what illness was the main cause for their inability.

Work ability score

The respondents were asked to compare their current work ability to their best lifetime work ability on a scale from 0 to 10, where a score of 0 represented full work disability and a score of 10 indicated work ability at its best. The question was presented to all those under 75 years of age who participated in the interview. It is part of the Work ability index (see below, Tuomi et al. 2006). The repeatability of the work ability score was excellent (Table 4.2.4, Chapter 4.2)

Work ability index

The work ability index (WAI) is based on a series of questions taking into consideration the physical and mental demands of work and the health and resources of the employee (Tuomi et al. 2006). It consists of seven items (six included in the interview):

1) Work ability score (see above): Current work ability compared with the lifetime best (interview question IB04)
2) Work ability in relation to the demands of the job (interview questions IB09 and IB10)
3) Number of current diseases diagnosed by a physician (interview, diseases BA-section)
4) Estimated work impairment due to diseases (interview question IB11 and IB12)
5) Sick leave during the past year (interview question IB13)
6) Own prognosis of work ability two years from now (interview question IB15)
7) Mental resources (Questionnaire 1, questions 49–51)

Some of the questions in the work ability index can be applied only to people in work life. Therefore, the index could only be used to determine work ability among those who had been working at any time during the previous 12 months (Gould et al. 2008).

Other questions on working ability

The other questions on work ability in the interview section IB concerned the past and future development of work ability, working while sick and factors that hinder respondents managing the work (those who were currently working) or taking part in working life (those not working). These factors included problems with health and work ability, lack of education or

skills, decreased motivation or desire to work, difficulties outside work and problems in working community/lack of jobs. The response options ranged from 1=not present or no harm to 3=hinders a lot. The respondents were also asked about their need for training and how they saw their prospects of re-employment if they were made redundant. All subjects except pensioners were asked whether they had considered retiring before the age for old-age pension fulltime or part-time.

Questionnaire 1 included a question whether any longstanding health problems hindered working, with a 10-point response scale. A corresponding question in Questionnaire 2 concerned musculoskeletal symptoms hindering working capacity. The repeatability of this question was good, α=0.83, but the measure showed statistically significantly lower value at the second assessment occasion (Table 4.2.1, Chapter 4.2).

10. HEALTH BEHAVIOURS

10.1 Smoking

Otto Ruokolainen

For all of the participants, a question: *"Have you ever smoked during your lifetime?"* (*"no", "yes"*) was presented in the interview (see Appendix 4). The rest of the questions on smoking were put to those who had smoked even once.

The next question was *"Have you ever smoked daily for at least one year?" ("no", "yes")*, and a follow-up question was presented to those who reported having smoked daily: *"How many years altogether have you smoked daily?"* (years, an open ended answer). The respondents' current smoking was enquired with a question: *"Do you smoke nowadays (cigarettes, cigars or pipe)?"*, and answer options were *"daily", "occasionally"* and *"not at all"*. An open-ended question *"On average, how much do you smoke or used to smoke daily (cigarettes, cigarillos, cigars, full pipes)?"* was presented to current smokers who had ever smoked daily for at least one year.

The last smoking occasion was asked with the question: *"When did you smoke last time?"*. The seven point answer scale for this question ranged from *"yesterday or today"* to *"over 10 years ago"*. A follow-up question with an open-ended answer was posed to those whose last smoking occasion was within the last ten years: *"How many times have you seriously (= been without smoking for 24 hours at a time) attempted to quit smoking during the past 10 years?"*.

In the phone interview, short questionnaire and questionnaire for young adults, only the question on current smoking was included.

10.2 Alcohol consumption

Janne Härkönen

The information concerning alcohol consumption was collected by Questionnaire 1 (see Appendix 5). The respondents were first asked to define, whether they were lifetime abstainers, had quit drinking (in what year), or were current drinkers. Guidelines how to measure a standard drink were given: one 330 ml standard bottle of medium strength beer, a small glass of wine or one shot of spirits. Alcohol consumption was measured with a beverage-specific quantity-frequency method. The frequency of each beverage type (beer, cider, alcopops; wine; spirits) was measured in ten answer choices, with six choices ranging from never to less frequently than weekly, and four choices covering one week's period. Quantity was covered by questions on the three beverage types. For mild beverages of beer and cider, nine choices were given, ranging from less than one bottle to 15 bottles or more. For wine, eight answer choices were given, ranging from less than one glass to two large bottles (à 0.75 litres) or more. For spirits, ten answer choices were used, ranging from one drink to more than two half-litre bottles. Beverage-specific quantities were first transformed into standard Finnish drinks (12 g of pure alcohol) and then multiplied by the frequency of the given beverage (for more information on measuring alcohol consumption, see e.g. Gmel and Rehm 2004).

Risky drinking was assessed using the AUDIT-C screen, which includes three items: the frequency of drinking, quantity of alcohol typically consumed, and the frequency of drinking six or more drinks on one occasion. Each AUDIT-C question has five answer choices, which are rated from zero to four, thus resulting in a total score of 0 – 12 points (scores of zero reflect no alcohol use in the past year). Following the Current Care Guidelines, a score of six or more was considered positive in men; in women, a score of five or more was considered positive (Käypahoitosuositus 2015).

AUDIT-C questions were also included in the questionnaire for young adults while only a single question on alcohol consumption (*"How often do you have a drink containing alcohol?"*) was included in the phone interview and short questionnaire.

10.3 Dietary habits

Satu Männistö, Annamari Lundqvist, Laura Sares-Jäske, Tuija Jääskeläinen, Noora Kanerva, Katri Sääksjärvi and Paul Knekt

There were several questions on dietary habits in the interview (see Appendix 4). Information on typical fat spread on bread and fat in cooking was collected by the standard questions (Helakorpi et al. 2000). Questions on consumption of vegetables, fruit and berries, rye bread and cheese were also presented. In addition, participants were asked whether they had been in nutritional counselling because of some illness during the past 5 years and who does mainly prepare their meals.

A question regarding consumption of vegetables was also included in the short questionnaire while no information on dietary habits was collected among those who only participated in the phone interview and among those young adults who only received a questionnaire.

Food Frequency Questionnaire (FFQ)

Information on food and nutrient intakes was collected by a food frequency questionnaire (FFQ). The FFQ has become the primary method in epidemiological studies concerned with the association of diet and the risk of diseases (Willett 2013) because it provides information on the individual's ordinary diet over a longer period. The FFQs are often designed to assess the individual's diet as a whole and the main aim is to rank individuals according to their food or nutrient intakes, not necessarily to measure the absolute intakes. The FFQ is easy for participants to complete and the answers can be easily computerized which makes it quite inexpensive to use. The development of the questionnaire itself, however, is a time-consuming exercise, and it is always necessary to ascertain the validity of the FFQ compared to food records or recalls.

The semi-quantitative FFQ was updated for the Health 2000 Survey from the questionnaire of the Kuopio Breast Cancer Study (Männistö et al. 1996, Paalanen et al. 2006). The content of the newest updated FFQ version for the Health 2011 Survey was based on the national FINDIET 2012 Study that included 48-h dietary recalls (Helldán et al. 2013). In general, the participants were asked to describe their usual diet over the past 12 months. The questionnaire listed 131 food items, mixed dishes and alcoholic beverages commonly used in Finland,

grouped in the following categories: dairy products; grain products; fat spreads; vegetables; potatoes, rice and pasta; meat; fish; poultry and eggs; fruit and berries; desserts; sweet and snacks; and beverages.

The average use of 131 foods was recorded by nine frequency categories ranging from never or seldom to at least six times a day. Participants could compare the size of their portions with predefined sizes printed on the questionnaire. The FFQ used in Health 2000 Survey had the same portion sizes for both sexes whereas in the newest FFQ some portion sizes were different for men and women based on the national FINDIET 2012 Study. The questionnaire also included separate questions on special diets and dietary supplements.

The FFQ was handed to all participants in the health examination and they were asked to complete it later at home. The questionnaire was introduced to each participant, and the filling instructions were reviewed together with them. A total 4,375 participants returned the FFQ.

Exclusions were made due to blank or incompletely filled FFQs (N=182) and daily energy intake cut-off points corresponding to 0.5 per cent at both ends of the daily energy intake distributions for men and women separately (N=42+42). Thus, in the end, intake of food and nutrients was calculated for 4,109 participants.

The average daily intakes of ingredient groups (e.g. rye, fish and berries), food groups (e.g. fish soups) and nutrients (e.g. energy-yielding nutrients, fibre and vitamin C) were calculated by using the National Food Composition Database (FINELI(R)) as well as the FINESSI (Reinivuo et al. 2010) software of THL. The final dietary dataset comprises around 100 ingredient groups, 100 food groups and 100 nutrients that can be used for research purposes.

The reproducibility of the FFQ versions has been measured twice (Männistö et al. 1996, Paalanen et al. 2006) and the validity compared with dietary records three times (Männistö et al. 1996, Paalanen et al. 2006, Kaartinen et al. 2011) over the last two decades. In those validation studies, the first evaluation of FFQ included diet as a whole, the second one concentrated more on the differences between sex, age and BMI, and the third one focused on carbohydrate fractions, dietary glycaemic index (GI) and the glycaemic load (GL). The reproducibility and validity results were similar compared to large internationally well-known studies (Pietinen et al. 1988, Willett 2013). As a consequence, the FFQ is reasonably accurate when the restrictions concerning some foods and nutrients are taken into account.

10.4 Physical activity

Tomi Mäki-Opas

Information on physical activity was collected in Questionnaire 1 (see Appendix 5). The instruction read as follows *"If there is major seasonal variation, select the option that best describes your average situation"*.

Leisure-time physical activity ("lifestyle") was assessed with two questions: *"How much do you exercise and strain yourself physically in your leisure time?"*. The response alternatives were: "*I mainly read, watch television, or do other activities that do not strain me physically", "I mainly walk, cycle, or move in other ways for at least 4 hours per week", "I do vigorous PA more than 3 hours per week"* and *"I participate regularly in competitive sports"*. Another question on leisure-time physical activity ("frequency") was *"How often do you exercise at least 30 minutes so that you sweat and get out of breath?"* The response alternatives were *"Daily", "4–6 times per week", "2– 3 times per week", "Once a week", "2–3 times per month"* and *"Few times per year or less frequently"*.

Commuting physical activity was assessed using the question *"How many minutes do you walk or cycle to and from work daily?"* The response alternatives were: *"I am not working or I work from home", "Use public transport or car during commuting", "Less than 15 minutes per day", "15–29 minutes per day", "30–59 minutes per day", "1–2 hours per day"* and *"Over 2 hours per day"*.

Occupational physical activity was inquired by posing the question *"How physically strenuous is your job physically?"* including response alternatives *"In my job I mainly sit and do not walk much", "I walk quite a bit in my job, but I do not need to lift or carry heavy items.", "In my job I need to walk or lift quite a lot or climb stairs or walk uphill"* and *"My job is heavy physical labour and I have to lift or carry heavy items, dig, shovel, pound or do some other heavy labour"*.

The questions on leisure-time physical activity ("lifestyle"), commuting physical activity and occupational physical activity have been shown to predict morbidity and mortality (Welin et al. 2003, Hu G et al. 2005). The leisure-time physical activity ("lifestyle") question has also been validated against accelerometer measurements (r=0.51, Fagt et al. 2011). The questions originate from the Gothenburg Study on Men Born in 1913 (Saltin and Grimby 1968, Grimby

et al. 1972, Wilhelmsen et al. 1973, Welin et al. 2003). All these three questions were also included in the Health 2000 Survey and quite similar questions were presented in the Mini-Finland Survey.

Health-enhancing physical activity was also inquired in the Health 2011 Survey. The questions on health-enhancing physical activity were developed by the UKK Institute and their aim was to measure whether people fulfil the current recommendations for health-enhancing physical activity (UKK Institute 2009, Husu et al. 2011). Information on average daily sitting time at work, at home, while travelling and elsewhere was also collected in Questionnaire 1.

Objective measurements of physical activity and fitness were collected in the sub-study on Physical Activity and Fitness (see Chapter 19.1).

10.5 Sleep and sleeping

Annamari Lundqvist

Several questions on sleep and sleeping were included in the self-administered Questionnaire 3. Sleep duration was measured by a single item asking how many hours the respondent sleeps on average during 24 hours. Sleep satisfaction was assessed by the question *"Do you think that you get enough sleep?"*. Respondents were also asked to report whether they snored and to ask others if they did not know. Those who responded *"yes"* were asked to answer more specific questions related to sleep apnoea. These questions concerned the frequency of snoring and loudness of snoring as well as respiratory arrests in sleeping. In addition, the participants were asked whether they had felt exceptional tiredness.

Some questions related to sleep and symptoms of insomnia were also presented in other parts of survey protocol. These include questions of e.g. seasonal variations (see Chapter 12) and health related quality of life (see Chapter 17).

11. HEALTH AND DISEASES

Seppo Koskinen, Annamari Lundqvist and Päivikki Koponen

Information on health and diseases was collected mainly in the interview (see Appendix 4) and Questionnaire 2 (see Appendix 5). In addition, a diagnostic mental health interview was carried out during the health examination (see Chapter 12). Key questions on health and diseases were also included in the phone interview and the short questionnaire as well as in the questionnaire for young adults (see Appendices 4 and 5).

The occurrence of major diseases was inquired in the interviews using a standard set of questions developed for the Mini-Finland Survey (Aromaa et al. 1989) and further elaborated for the Health 2000 Survey (Reunanen and Heliövaara 2008). If feasible and available, international standard questions were used and specified below for each topic. Questions on the treatment of diseases and use of medicines were organized into separate sections in the interview and they are described in Chapters 18.1 and 18.2, respectively. Methods for assessing mental health and oral health are presented in Chapters 12 and 13, respectively.

11.1 Self-rated health and long-standing illnesses

Information on perceived health was obtained in the interview by a standard question with five preset responses from good to poor (de Bruin et al. 1996). The respondents were also asked whether they had any permanent or chronic illness or injury which reduces their working capacity or functional ability. Further questions on changes of job or working tasks due to illnesses during the past decade were presented to those who had been employed at any time during the last 10 years.

11.2 Cardiovascular diseases and diabetes

The interview included structured questions concerning heart diseases (myocardial infarction, coronary heart disease (angina pectoris), heart failure, arrhythmia, congenital heart disease, valvular heart disease, heart muscle disease (cardiomyopathia), inflammation of the heart muscle (myocarditis) and any other heart disease). Other cardiovascular diseases covered by the interview were hypertension, stroke and varicose veins in the lower extremities. The subjects were first asked whether a doctor had ever diagnosed them with any of the diseases specified above and if so, whether they were currently receiving medication and/or seeing a doctor or a public health nurse because of it. Subjects with myocardial infarction or stroke were also asked whether they had been in hospital because of the disease. For subjects with coronary heart disease, it was also recorded how much coronary disease symptoms limit their daily activities using the four-class NYHA classification (The Criteria Committee of the New York Heart Association, 1994) and whether they had had an invasive coronary procedure (bypass-surgery or balloon distension). Persons reporting myocardial infarction or stroke were asked when the disease had been diagnosed for the first time. Among subjects with high blood pressure, the number of blood pressure measurements by health care professionals during the past 12 months was recorded. Further, persons reporting varicose veins were asked whether they had been operated.

Exertion-associated chest pain symptoms indicative of coronary heart disease, intense attacks of chest pain and symptoms of intermittent claudication indicating arterial occlusion of the lower extremities were queried in the self-administered Questionnaire 2 with a set of questions originally developed by Geoffrey Rose and recommended by the WHO for use in population studies (Rose and Blackburn 1968, Rose et al. 1982). Questionnaire 2 also included questions on shortness of breath upon physical exertion as described in Chapter 11.4.

Information on diabetes was collected in the health interview. For those reporting that a doctor had diagnosed them with diabetes, further questions on time of diagnosis, treatment, blood glucose measurements and complications (retinopathy, albuminuria, nephropathy, amputation of a part of the lower limb) were presented.

In addition, blood lipids and glucose were analysed among those who participated in the health examination as has been illustrated in Chapter 5. Further, blood pressure and resting ECG measurements were also included in health examination (see Chapter 7).

11.3 Musculoskeletal disorders and injuries due to the accidents

The interview included questions on the following musculoskeletal disorders diagnosed by a physician: rheumatoid arthritis, osteoarthritis (arthrosis, joint degeneration), osteoporosis, neck disease, back disease and fractures. For those who reported any of the above conditions, further questions regarding visits to a doctor or public health nurse, operations, use of physiotherapy, and use of medication because of the disease were presented. In the case of arthritic conditions, neck disease and back disease, the specific type of disease was also recorded. For osteoporosis and bone fractures, the site of condition was recorded.

Information on musculoskeletal symptoms was collected by the self-administered Questionnaire 2. Subjects were first asked whether they had experienced back pain, neck pain or shoulder pain. Among those reporting pain, more detailed information about the symptoms was collected. In addition, the questionnaire included a question concerning pain, ache or motion sensitivity in one or more joints during the past 30 days. Persons reporting joint symptoms were asked to identify the joints with symptoms in a diagram of the human body. Further, all participants were asked whether they had been unable to perform their daily tasks due to joint pain within the past 30 days and to estimate on a scale from 1 (no at all) to 10 (worst possible) how much hindrance they perceived at work and during leisure time. Information on walking difficulties and falls was also collected.

To assess physical functional ability, the health examination included several tests described in Chapter 14.

Information on permanent injuries or defects caused by an accident was collected in the interview (see Appendix 4).

11.4 Respiratory diseases

In the interview, the subjects were asked whether a doctor had ever diagnosed them with asthma, chronic obstructive pulmonary disease (COPD) or chronic bronchitis. Those who reported any of the above diseases were further asked whether they were currently receiving medication and/or seeing doctor or public health nurse because of the disease. For those reporting asthma further questions on shortness of breath were presented.

Shortness of breath upon physical exertion was queried with four questions included in Questionnaire 2. Those questions are in line with the recommendations of the British MRC (Fletcher et al. 1959) and the WHO (Rose and Blackburn 1968). As shortness of breath in connection with physical exertion may be due to respiratory disease, circulatory diseases or poor physical fitness, the purpose of the questions was simply to determine the presence of symptom rather than to establish the underlying cause of the symptom.

Furthermore, one question on breathing was included in the quality of life index (15D) which is described in Chapter 17. As a part of the health examination, spirometry measurements were taken (see Chapter 7.3).

11.5 Skin diseases and allergies

The interview included a question on any longstanding dermatological (skin) disease diagnosed by a doctor. Those giving a positive answer were asked to specify whether the condition was allergic eczema, atopic eczema, toxic eczema, fungal infection (in skin or nails), psoriasis or other skin disease. In addition, the self-administered Questionnaire 2 included questions on allergic rhinitis, allergic eye inflammation and atopic eczema. Information on both symptoms and diseases diagnosed by a doctor was collected.

11.6 Reproductive and sexual health

All questions on reproductive health were based on the questions included in previous Finnish surveys (no international standard instruments were available). Most of these questions had also been used in the Health 2000 Survey, but a few questions were adapted and specified, e.g. due to changes in clinical treatment and recommended contraceptive methods, or based on analyses of the Health 2000 data. In the Health 2011 Survey, the interview included questions for both men and women about the number of children they had. Both men and women younger than 70 years were also asked about experiences of infertility, seeking examinations and treatments due to infertility, and the success of these treatments (getting a child). Both men and women younger than 55 years were asked about contraception. Women also received questions on gynaecologist examinations and breast self-examinations. Women below the age of 70 years were asked about menstruation and menopause, as well as hysterectomy and

hormone replacement therapy. Women below 55 years of age were asked about the number of induced abortions, miscarriages and childbirths, having ever had pregnancy complications (high blood pressure, preeclampsia and gestational diabetes) as well as duration of breastfeeding.

The self-administered Questionnaire 1 for young adults comprised additional questions on satisfaction with sex life, the frequency of sexual interaction, the number of partners during the past 12 months, and whether they had used a condom in case they had had sexual intercourse with someone else than their permanent partner within the past 12 months. The postal questionnaire for young adults included, for both men and women, questions on contraception. They were also asked if they had used a condom in case they had had sexual intercourse with someone else than their permanent partner within the past 12 months. At the end of this questionnaire there were additional questions for women about the number of induced abortions, miscarriages and childbirths.

Questions on screening and health examinations also comprised items related to reproductive and sexual health (e.g. mammography and cervical cancer screening, see Chapter 19).

11.7 Other diseases

Other diseases covered by the general health interview were cataract, glaucoma, eye ground degeneration, other visual defect, hearing defect, bowel disease, cancer, benign tumours of the uterus (e.g. myoma), Parkinson's disease, urinary tract infection, urinary incontinence and headache. The subjects were first asked whether a doctor had diagnosed them with the disease, and those giving a positive answer received a few further questions on e.g. the location, treatment and symptoms of the disease. The subsequent questions varied from one disease to another.

12. MENTAL HEALTH

Jaana Suvisaari, Jens Strehle, Olavi Lindfors, Timo Partonen, Satu Viertiö, Olli Kiviruusu, Sami Pirkola, Aino Mattila and Kirsi Ahola

Mental health and substance use-related problems were assessed with questionnaires and with a structured psychiatric interview, which covered depressive, anxiety and alcohol use disorders and psychotic symptoms. Questionnaires assessed mood and anxiety symptoms, psychological distress and burnout. In addition, symptoms related to eating disorders were asked from young adults. The general interview included questions on psychiatric and substance use disorders diagnosed by a physician, and health care contacts during the past 12 months for these disorders.

Information on factors related to mental health and well-being were collected in three different ways:

1) Assessments based primarily on the structured psychiatric interview M-CIDI (mood disorders, alcohol use disorders, psychotic symptoms and anxiety disorders)

2) Assessments based primarily on questionnaires (Beck Depression Inventory (BDI-13); General Health Questionnaire (GHQ-12); Toronto Alexithymia Scale (TAS-20); Hopkins Symptom Checklist-25 (HSCL-25); Maslach Burnout Inventory-General Survey (MBI-GS); Seasonal Pattern Assessment Questionnaire (SPAQ); an eating disorders screen (SCOFF) for people aged 18–29 years)

3) Assessments based on the general health interview (self-reported treatment contacts, perceived need for treatment and quality of care received)

The location of instruments used for the assessment of mental health in the study questionnaires and interviews is presented in Table 12.0.1.

Table 12.0.1. The position of instruments used for the assessment of mental health.

	Health examination	Questionnaire 1	Questionnaire 2	Questionnaire 3
CIDI	x			
Treatment	x			
BDI-13			x	
GHQ-12		x		
TAS-20				x
HSCL-25				x
MBI-GS		x		
SPAQ				x
SCOFF		x		

CIDI = Composite International Diagnostic Interview; BDI-13 = Beck Depression Inventory; GHQ-12 = General Health Questionnaire; TAS-20 = Toronto Alexithymia Scale; HSCL-25 = Hopkins Symptom Checklist-25; MBI-GS = Maslach Burnout Inventory-General Survey; SPAG = Seasonal Pattern Assessment Questionnaire; SCOFF = an eating disorders screen

12.1 Mental health interview

The interview method chosen for the study of mental health disorders in the Health 2000 Survey was the Munich-Composite International Diagnostic Interview or M-CIDI (Wittchen et al. 1998). CIDI is a structured interview developed by the World Health Organization (WHO) in 1990 for the purpose of epidemiological research. Some interviewer training is needed, but the interviewers do not need to be mental health professionals. The method is designed for use in different cultures. Two different versions of the CIDI interview are available: the Munich-Composite International Diagnostic Interview (M-CIDI, Wittchen et al. 1998) and WHO CIDI (Kessler and Ustun 2004). At the time of the Health 2000 Survey, only the M-CIDI was available as a computerized interview and therefore it was chosen for the study. Both CIDI versions have been used extensively in general population mental health surveys (Jacobi et al. 2004, Kessler et al. 2007).

A key advantage of the CIDI over other interview methods lies in its extensive international use and its efficient training and translation organisation. Several reliability and validity studies have been published on the method (Wittchen et al. 1998). Its main drawback is that the

interview lasts on average 75 minutes, and therefore it could not be implemented in full in the Health 2000 Survey. The Health 2000 CIDI interview covered only the most common mental disorders: mood, anxiety and substance use disorders and psychotic symptoms. Since the time reserved for the CIDI interview was even shorter in the Health 2011 Survey, two sections (assessing manic symptoms and other substance use disorders than alcohol) were omitted from the Health 2011 CIDI interview. The number of questions was also reduced so that only questions essential for the diagnostic algorithms were retained. The CIDI version used in the Health 2011 Survey covered eight diagnoses: panic disorder, agoraphobia, social phobia, generalized anxiety disorder, dysthymia and major depressive disorder as well as alcohol abuse and dependence. The interview covered symptoms and signs present during the past 12 months. For alcohol use disorders, lifetime disorders were also assessed.

The development of CIDI version used in Health 2000 Survey is described in detail elsewhere (Pirkola et al. 2008). Before the Health 2011 Survey started, the CIDI translation was checked again against the original German version and the English translation, and linguistic errors were corrected. The new CIDI version was programmed by Jens Strehle, M.Sc., from the University of Dresden. The new CIDI version was piloted in May 2011, and all errors that were observed then were corrected before the fieldwork.

12.2 Mental health questionnaires

Current psychological distress was assessed using the 12-item version of the General Health Questionnaire (GHQ-12, Goldberg 1972). GHQ-12 includes 12 questions assessing symptoms commonly related to depression as well as general functioning, e.g. ability to face problems and make decisions. All items have a 4-point scoring system ranging from a "*better/healthier than normal"* option, through a "*same as usual"* and a "*worse/more than usual*" to a "*much worse/more than usual"* option. These are scored using a 0-0-1-1- scoring, so that *"better"* and *"usual"* responses are scored as 0, and *"worse"* and *"much worse"* responses are scored as 1. The responses to individual items are added to give a total score which varies from 0 to 12.

Current depressive symptoms were screened using the 13-item version of the Beck Depression Inventory (BDI-13, Beck and Beck 1972). In the Health 2000 Survey, the 21-item version of the BDI was used, but since the 13-item version has been shown to perform equally well (Aalto et al. 2012), the shorter version was selected for the Health 2011 Survey. The scoring of each of the BDI-13 questions is described in Table 12.2.1. The scores of individual items are added

to give the total score which varies from 0 to 39. In addition, some methods assessing psychological functioning are described in Chapter 14.3

Table 12.2.1. Scoring of the Beck Depression inventory-13 (BDI-13).

Question 1:	1=0	2=1	3=2	4=2	5=3
Question 2:	1=0	2=1	3=2	4=2	5=3
Question 3:	1=0	2=1	3=2	4=2	5=3
Question 4:	1=0	2=1	3=2	4=2	5=3
Question 5:	1=0	2=1	3=2	4=2	5=3
Question 6:	1=0	2=1	3=2	4=2	5=3
Question 7:	1=0	2=1	3=2	4=2	5=3
Question 8:	1=0	2=1	3=2	4=3	
Question 9:	1=0	2=1	3=2	4=3	
Question 10:	1=0	2=1	3=2	4=3	
Question 11:	1=0	2=1	3=2	4=2	5=3
Question 12:	1=0	2=1	3=2	4=3	
Question 13:	1=0	2=1	3=2	4=3	

As a new questionnaire, Hopkins Symptom Checklist-25 (HSCL-25), a short version of the SCL-90 instrument (Derogatis et al. 1974), was included to assess current depressive and anxiety symptoms. It has been used in several previous Finnish studies to screen for mental disorders (Joukamaa et al. 1994, Veijola et al. 2003). In HSCL-25, participants are asked to report the severity of each symptom asked in the questionnaire on a 4-point scale, ranging from 1 (not at all) to 4 (extremely). Responses are summed and divided by the number of answered items to generate a total score. A total score exceeding 1.75 was used to indicate severe depressive and anxiety symptoms. Separate scores based on anxiety symptoms (10 items) and depressive symptoms (15 items), can be formed.

Burnout was measured with the Maslach Burnout Inventory–General Survey (MBI-GS, Maslach et al. 1996, Kalimo et al. 2006). The MBI-GS consists of three subscales: exhaustion,

cynicism, and professional efficacy. The items were scored on a 7-point frequency rating scale ranging from 0=never to 6=daily. High scores on exhaustion and cynicism, and low scores on professional efficacy, are indicative of burnout. The items of professional efficacy were reversed (lack of professional efficacy). One missing value per subscale was allowed. A weighted averaged summary score of the dimensional scores was calculated (0.4 x exhaustion + 0.3 x depersonalization + 0.3 x lack of personal accomplishment). The summary score was categorized as no burnout (0-1.49), mild burnout (1.50-3.49), and severe burnout (3.50-6).

The 20-item Toronto Alexithymia Scale (TAS-20, Bagby et al. 1994) was used to measure alexithymia, which refers to a personality construct characterized by impoverishment of fantasy, poor capacity for symbolic thought, and difficulties in experiencing and verbalizing emotions. TAS-20 includes 20 questions related to difficulties in identifying feelings, difficulties in describing feelings, and externally oriented thinking. The items in the TAS-20 are statements on a 5-point Likert scale ranging from strongly disagree to strongly agree. Five items (items 4, 5, 10, 18 and 19) are reverse coded. The minimum total score of the TAS-20 is 20 and the maximum is 100 (Mattila et al. 2006).

Seasonal variations in mood and behaviour were assessed with seven items derived and adapted from the Seasonal Pattern Assessment Questionnaire (SPAQ, Rosenthal et al. 1984), including the six seasonal variations in sleep duration, social activity, mood, weight, appetite, and energy level. Two modifications were made to the original scoring as follows. Each item was scored from zero to three (none, slight, moderate or marked), not from zero to four, with the sum or global seasonality score (GSS) ranging from 0 to 18. The psychometric properties of this modified questionnaire have been tested and shown to be good (Rintamäki et al. 2008). The 7/8 cut-off score was applied for the two GSS categories (0-7 vs. 8-18). In addition, there was a question: *"If you experience changes by seasons, do you feel that these are a problem for you?*" This item was scored from zero to one (no variations, variations but no problem) and two to five (variations of mild, moderate, marked or severe problem), not from zero to five.

SCOFF is a five item questionnaire assessing symptoms related to eating disorders. The acronym SCOFF is created from the questions: 1) Do you make yourself Sick because you feel uncomfortably full? 2) Do you worry you have lost control over how much you eat? 3) Have you recently lost more than one stone (6 kg in the Finnish version) in a 3 month period? 4) Do you believe yourself to be fat when others say you are too thin? 5) Would you say that food dominates your life? A threshold of two positive answers has been proposed to raise a suspicion of an existing eating disorder (Morgan et al. 1999, Lähteenmäki et al. 2009).

13. ORAL HEALTH

Miira Vehkalahti and Liisa Suominen

Information on oral health collected in the interview and questionnaires covered the whole sample whereas the clinical oral health examinations were carried out in two out of the five field examination areas, i.e. the southern and northern (Helsinki and Oulu), and the radiographic examination in Helsinki location only. The protocol for the clinical oral health examinations was slightly modified from that used in the Health 2000 Survey (Suominen-Taipale et al. 2008) by going more into details and adding some measurements, but still ensuring the comparability with the previous data. The more detailed measurements covered dental plaque and gingival bleeding by tooth. The new measurements involved the presence of fixed prosthetics and the material of restorations. The clinical examination ended by two new questions about temporomandibular symptoms and four questions about the use of removable dentures.

13.1 Oral health in the interview and questionnaires

The oral health-related questions were largely the same as in the Health 2000 Survey (Suominen-Taipale et al. 2008) and in the Mini-Finland Study (Vehkalahti et al. 1991). The questions in the interview covered self-reported oral health status, oral self-care, use of services and experiences as an oral health care customer. In Questionnaire 1, information on eating sweets and drinking sweetened drinks was collected by asking how often the participants add sugar to tea or coffee, drink other drinks with sugar added (e.g. juices, lemonades, hot chocolate), or use chewing gum with xylitol. The response options ranged from never to three times per day or more often. Information on oral health-related problems was collected by using the 14-item Oral Health Impact Profile (OHIP-14, Slade and Spencer 1994, Slade 1997), included in Questionnaire 3. Some minor adaptations were made in the questions on oral health services due to changes in the oral health service system.

13.2 Clinical examination of the mouth

The oral health examination team comprised a dentist and a dental nurse. Overall the examination took 15 minutes. Following the clinical examination the dentist offered the participants a written summary of the clinical and radiographic observations whenever taken with a recommendation for seeking oral health care if necessary.

Examinations were carried out with a portable dental treatment unit (Dentronic Mini-Dent®, Planmeca Oy) comprising a built-in compressor, saliva suction and a high-powered suction engine. In addition, the equipment comprised a portable patient chair, fibre optic light (Novar®), fibre optic head lamp (Tekmala Oy) and a letter scale. The team had at hand written detailed instructions on the stages of the clinical examination, measurement determinations and for making computer entries. The participants were first asked*: "Do you have any health condition for which your doctor or dentist has said you require antibiotic protection in connection with oral health care?"* If yes, periodontal measurements were excluded.

The clinical examinations always followed the same order. First, maximum mouth opening was measured and the participant's jaw joints and masticatory muscles were palpated and the related sounds and pain were observed. Information was recorded on the presence of removable dentures, their condition, fit and cleanliness. After the examination of the oral mucosa and occluding tooth pairs, the patient chair was adjusted to the reclined position. The next steps were identification of teeth, recording of the presence of dental plaque excluding wisdom teeth and assessment of spaces in the dental arches excluding the molar areas. Tooth-based recordings included the status of all teeth and the material of restorations as well as depth of periodontal pockets and bleeding on probing, the two latter excluding wisdom teeth. The presence of fixed prostheses was recorded separately for each sextant.

The participants' teeth were always examined in the same order, starting from the last tooth in the upper right quadrant and ending with the last tooth in the lower right quadrant. The identification of a tooth and the determination of the status of teeth and periodontium were based on the methodology used in the Health 2000 survey (Suominen-Taipale et al. 2008) and in the Mini-Finland Survey (Vehkalahti et al. 1991). All tooth surfaces were examined and the observations were combined and recorded by tooth. Details of the tooth-based clinical recordings are shown in Table 13.2.1.

Finally, the participant was asked about pain in temporomandibular region and about possible symptoms when opening the mouth. In case of removable dentures, the participants were

asked whether any repairs had been made or needed now, and to report the actual use of the dentures and about the occurrence of denture-related problems while eating.

Table 13.2.1. Description of tooth-based clinical recordings.

Measurement	Recordings
Identification of each tooth (32 teeth)	present; absent
Presence of dental plaque (28 teeth), buccal surfaces	no plaque; any plaque
Dental status (32 teeth)	sound; fractured; filled; coronal caries; root caries; carious radix; non-carious radix
For each tooth with a restoration: material of filling	amalgam; composite; temporary; other; prosthetic crown
Depth of periodontal pockets (28 teeth)	no pocket; 4<6 mm; 6+ mm pocket
Presence of gingival bleeding on probing (28 teeth)	present; absent

13.3 Radiographic examination

In Helsinki, the field team had access to a digital panoramic x-ray unit (Promax 2D®, Planmeca Oyj) as well as a PC with software (Romexis®, Planmeca Oyj) for the preliminary examination of the radiographs. The imaging values were adjusted according to the participant's size and varied within the range of 60-68 kV and 8-13 mA. At the examination site, the equipment was installed and calibrated by an expert from the Planmeca Oyj. The images were stored in electronic format and a personal copy was handed to the participant.

13.4 Quality assurance of clinical and radiological measurements

Clinical measurements were based on validated methods used in previous oral health population surveys. Measurement determinations were designed with a specific view to clarity and unambiguousness. The piloting of the clinical examination, the use of electronic

forms, induction training to field examination personnel and the detailed guidelines provided for the field teams were the imperative means to assure measurement quality.

A four-day pilot study of the clinical measurements took place at a public dental clinic of the Helsinki City Health Department between the 1 and 7 April 2011. The examining dentist and dental nurse were the same as in the piloting for the Health 2000 Survey. A total of 32 patients were examined under supervision of the Health 2011 oral health planning team.

The training course for the oral health examination included lectures and hands-on-sessions where the fieldwork personnel received detailed instructions covering all the clinical and radiological measurements and their recordings. Each examiner also received a protocol with written and illustrated definitions for the measurements. The training course ended with a dry run in real circumstances.

During the field examinations, the dentist who took part in pilot acted as a reference examiner and carried out parallel clinical examinations on a total of 75 participants. Similarly, members of the planning team made several visits to the field locations in order to control the procedures and the quality of the clinical examinations.

14. FUNCTIONING

Päivi Sainio

WHO's International Classification of Functioning, Disability and Health (ICF, WHO 2001) is a biopsychosocial approach to describe and structure functioning and disability. It portrays functioning and disability as a dynamic interaction between the health conditions and personal factors of the individual and the contextual factors of the environment. The ICF framework classifies human functioning on three levels: functioning at the level of body or body part, the whole person (activities), and the whole person in a social context (participation, WHO 2002). The ICF provides a framework for the definition and operationalization of disability also in population surveys, and it has been accepted as a framework for developing functioning and disability measures for population surveys by e.g. Eurostat, WHO and Washington Group on Disability statistics. Although the Health 2011 Survey was not originally designed on the basis of the ICF framework, the topics of functioning and methods to measure them quite comprehensively cover the various components of ICF. The Health 2011 Survey included e.g. physical, psychological, cognitive and social functioning, work ability and usual and basic activities of daily living. Information was also gathered on various environmental and personal factors as well as on the health conditions. The majority of the methods used to measure functioning in the Health 2011 Survey are described in this chapter, but some of them are described in other sections of this report as they can be used within other frameworks as well (Table 14.0.1). The description of the methods used already in the Health 2000 Survey is based on descriptions included in the *"Methodology Report. Health 2000 Survey"* (Aromaa 2008, Era and Sainio 2008, Rudanko and Koskinen 2008, Sainio and Malmberg 2008, Suutama et al. 2008).

Table 14.0.1. The position of the method description in this report by the measured construct.

Construct	Chapter
Physical functioning	Chapter 14.1
Physical fitness	Chapter 19.1
Lung and heart functions	Chapter 7
Vision, hearing	Chapter 14.2
Cognitive functioning	Chapter 14.4
Psychological functioning	Chapters 14.3 and 12
Social functioning	Chapter 14.5
Social capital	Chapter 16
Basic and usual activities (ADL, IADL)	Chapter 14.6
Quality of life	Chapter 17
Work ability and working conditions	Chapter 9
Living environment and housing	Chapter 8.4
Use of assistive devices, use and need of help	Chapter 15

14.1 Physical functioning

Päivi Sainio, Sari Stenholm and Markku Heliövaara

The assessment of physical functioning was based on self-report and performance tests (Table 14.1.1). The methods are well-established and widely used in population surveys and clinical studies (McWhinnie 1981, Guralnik et al. 1994, Guralnik et al. 1995a, Curb et al. 2006).

Table 14.1.1. The position of methods for the assessment of physical functioning in the Health 2011 Survey protocol.

Indicator	Interview	Health examination	Questionnaires	Physical Activity and Fitness[1)]	COURAGE[2)]
Questions on mobility (walking, running, stair climbing etc.)	HB02-HB11		Q1: Kys1_K0101 Q2: OIRE_59, OIRE_60 Q3: Kys3_LIIK		x
Walk test (4m/6,1m)		x			x
Chair stand (5/10 times)		x			
Hand grip strength		x (Jamar/Saehan-dynamometer)			x (Smedley-dynamometer)
Balance test		x (Guralnik's test, 70+ yrs)		x (one leg stand, 30–74 yrs)	
Squatting		x (55+ yrs)			
Movements of shoulder		x (55+ yrs)			
Vertical jump				x	
Modified push-ups				x	
6-minute walk test				x	
Heart and lung function[3)]		x			

Q1 = Questionnaire 1; Q2 = Questionnaire 2; Q3 = Questionnaire 3
[1)] Described in Chapter 19.1
[2)] Described in Chapter 19.2
[3)] Described in Chapter 7

Self-reported physical functioning

Assessment of self-reported physical functioning focused on mobility. The first question in the interview was on general ability to move *"How well can you move about?"*, followed by questions about ability to run a longer distance (about 0.5 km) or a shorter distance (about 100 meters), ability to climb up several flights and one flight of stairs, ability to walk 2 km and 0.5 km and to move about from one room to another, and ability to travel by train, bus or tram

(Rosow and Breslau 1966, McWhinnie 1981, Aromaa et al. 1989, see Appendix 4: HB-section and Table 14.1.1). The core question was *"Are you able to ...?"* The response options were: without difficulties, with minor difficulties, with major difficulties, not at all. To identify persons at a very early stage of mobility limitation (preclinical mobility limitation), three additional questions were posed to participants reporting no difficulty in climbing up several flights of stairs and in walking 2 km: *"How easy is it for you to [climb up several flights of stairs / walk 2 km without resting]?"*, *"Do you get tired when [climbing up several flights of stairs / walking 2 km without resting]?"* and *"During the past year have you cut down on [climbing up several flights of stairs / walking 2 km without resting] because your physical condition or health has deteriorated?"* (Fried et al. 1991, Simonsick et al. 2001, Mänty et al. 2007). In addition to above mentioned questions on mobility in the interview, one mobility item was included quality of life measures (both the EQ-5D and the 15D, see Chapter 17). Two questions about difficulties in walking or limping due to hip or knee problems were asked in Questionnaire 2 (see Appendix 5 and Table 14.1.1)

Repeatability of the prevalence of difficulties in walking 2 km was good (kappa 0.86). For the prevalence of limping/walking difficulties due to hip or knee problems, the repeatability was somewhat lower (kappa 0.73–0.76, see Table 4.2.2 in Chapter 4.2).

The questionnaire for young adults included only one question about mobility: running 0.5 km. A smaller sample of young adults was invited to the health examination, in which the more strenuous items of mobility were asked (running 0.5 km / 100 m, stair climbing several flights, walking 2 km, see Appendices 4 and 5).

Performance tests for physical functioning

Several performance tests measuring mobility, strength, balance and joint functions were conducted. The tests were performed in a standard order (see Figure 2.4.1). Before each test the study nurse ensured that the test could be safely administered. If the test was not conducted (e.g. due to severe disability), it was recorded with the reason in the data collection program. The study nurse explained and showed the test movements to the subject prior to each test.

The walk test

The walk test was conducted for subjects who were able to walk without help of another person.

The walk test was conducted over a distance of 4 m and 6.1 meters (Fiatarone et al. 1994, Guralnik et al. 1995b). Start line, 4 m line and end line (6.1 m) were marked on the corridor floor using tape. The subjects were asked to walk the 6.1 m distance as quickly as they could, starting from their normal, standing posture behind the start line and continuing at full speed beyond the end line. The subjects were allowed to use a walking aid if they normally used it. The stopwatch was started at the beginning of the test, split time was taken when the subject passed the 4 m line, and the timing was stopped when the subject crossed the end line. The study nurse stood at the 4 m line where it could easily be seen when the subject passed the 4 m and 6.1 m lines. If needed for safety reasons, the nurse walked with the subject (but was not allowed to help them). The times for walking 4 m and 6.1 m distances were recorded in the data collection program, which converted them to speed (m/s). In the home health examination walking speed was measured over 4 m distance if the necessary space for the longer distance was not available.

The repeatability of walking speed between observers was moderate (α=0.81, see Table 4.2.4 in Chapter 4.2).

The chair stand test

The chair stand test (Csuka and McCarty 1985, Guralnik et al. 1994) was conducted with a standard chair with no arm rests, a seat height of 43–45 cm from the floor and seat depth 39–43 cm. The back of the chair was placed against a firm table or wall. The subjects were asked to sit on the chair, with arms crossed in front of the chest and feet on the floor and slightly apart. From this position, they were asked to stand up once. If this did not succeed or hands were used to support the rising, the test was ended and the performance recorded. If the subjects managed to get up without using hands, they were asked to get up and sit down 10 times as quickly as possible. A split time was taken at five stands and timekeeping was ended after 10 stands. Both times were recorded in the data collection program. If the subject was not able to stand 10 times, the number of stands was also recorded. The test was discontinued if it was not completed in 90 seconds, or if it posed any risk to the subject's safety. In the home examination, a normal chair (not sofa) in the subject's home was used.

The repeatability of chair stand time was fair (α=0.71-0.77, see Table 4.2.4 in Chapter 4.2). More information on the reliability and validity of the test can be found in the TOIMIA-database (in Finnish): http://www.thl.fi/toimia/tietokanta/mittariversio/155/.

The joint functions tests

Four tests were used to measure joint functions, one designed for the lower limbs in general (squatting) and three for mobility of the humeroscapular joints. The tests were conducted for persons aged 55 years or older. Two of the tests (squatting and abduction of the upper arms) were included also in the Health 2000 Survey (Sainio and Malmberg 2008).

Squatting. The subjects were asked to squat and stand up. Light support for balance from a table edge was allowed. Performance was rated as 1) normal, if squatting down (thighs at horizontal level, or thighs and calfs touching) and getting up without support was successful; 2) restricted, if some support was needed or if the squat was not full (thighs not reaching horizontal level); 3) failed, if the subject could not stand up without using notable support or knees were flexed less than 45 degrees (Sievers et al. 1985).

Abduction of the upper arms. The subjects were asked to abduct both arms towards ceiling. Each arm was rated separately: 1) normal, if the arm was raised up (near the head, 30 degrees short of vertical line was accepted); 2) restricted, if the abduction was above horizontal level but not all the way up; 3) failed, if the abduction remained below horizontal level (Sievers et al. 1985).

External rotation of the shoulder joints. The subjects were asked to raise one arm at a time behind the head to reach with fingers the upper corner of the opposite scapula. The performance was rated as 1) normal, if scapula was reached; 2) restricted, if only the 7th cervical vertebra was reached; 3) failed, if the movement was less (Hoppenfeld 1976).

Internal rotation of the shoulder joints. The subjects were asked to rotate one arm at a time behind the back to reach with fingers the lower corner of the opposite scapula. The performance was rated as 1) normal, if scapula was reached; 2) restricted, if the fingers reached only the waist level; 3) failed, if the movement was less (Hoppenfeld 1976).

Repeatability of joint functions varied from low to fair: squatting kappa 0.60; abduction kappa 0.55–0.65; external rotation kappa 0.38–0.41; internal rotation kappa 0.70–0.75 (see Table 4.2.3 in Chapter 4.2).

The standing balance test

The standing balance test was conducted for subjects aged 70 years or older who could stand by themselves. Balance was measured using a progressive standing test (called Guralnik's test, Guralnik et al. 1994). First, the subjects were asked to stand still in a semi-tandem position (one foot in front of the other so that the side of the big toe of the trailing foot touched the heel of the leading foot) for 10 seconds. If successful, the subjects were asked to stand in a more difficult tandem position, with feet on the same line one directly behind the other, for 10 seconds. If the semi-tandem position failed, the subjects were asked to stand in an easier position, with feet side by side, touching each other for 10 seconds. The performance for each position was recorded as succeeded (10 seconds) or failed. If failed, the time in the position was recorded.

Hand grip strength

The hand grip strength was measured with Jamar/Saehan dynamometer (Sammong Preston Rolyan 2003) from the dominating hand, which was defined as the writing hand. If the subject could not use the dominating hand due to severe injury or disorder (e.g. cast due to fracture, hemiplegia), the measurement was conducted with the non-dominating hand. The size of the grip handle was adjusted according to the size of the subject's hand. The width of the grip was appropriate, when the middle joint of the index finger was in a 90 degree angle. The subjects were asked whether they felt comfortable with the width of the grip. The subjects sat straight in a chair, feet slightly apart on the floor. They held the dynamometer with wrist in a neutral position (i.e. in slight dorsal flexion) and elbow in 90 degrees. The opposite upper limb was resting on the lap. The subjects were asked to grip the handle as hard as they could for 3–5 seconds; throughout this time the study nurse urged the subjects to do their best. The second measurement was conducted 30 seconds later. If the difference between the two measurements was greater than 10 per cent, a third measurement was conducted again 30 seconds later.

14.2 Vision and hearing

Päivi Sainio

Vision was assessed on the basis of self-report and with vision charts and hearing on the basis of self-report only (Table 14.2.1).

Table 14.2.1. The position of methods used to assess vision and hearing.

Indicator	Interview	Health examination	Questionnaire 3	COURAGE[1)]
Near vision	HC01	Visual acuity test	Kys3_NAKO	x[2)]
Distant vision		Visual acuity test	Kys3_NAKO	x[2)]
Hearing	HC04, HC06		Kys3_KUUL	x[3)]

[1)] Described in Chapter 19.2
[2)] Test and questions
[3)] Questions

Self-reported vision and hearing

Vision was assessed in the interview by the question: *"Are you able to read newspaper text?"*, with three response categories (no difficulties, difficulties, not able). In Questionnaire 3, one question on near and distant vision was included in the 15D quality of life measure, with five response categories from normal vision to almost complete blindness. Hearing was assessed in the interview by the question *"Can you hear what is said in a conversation with several persons?"*, and *"Can you hear what is said in a conversation between two persons"*. Response categories were no difficulties, difficulties, and not able. In Questionnaire 3, one question on hearing was included in the 15D, with five response categories from normal hearing to completely deaf (see Chapter 17).

Questions about near vision and hearing were presented also to the young adults (see Appendix 4).

Visual acuity tests

Binocular visual acuity was measured using well illuminated distant and near vision charts (Oriola). Eye glasses or contact lenses were allowed if normally worn by the subject. Illumination was adjusted using the lights at each examination site and an additional spotlight. The test was not performed, if the subject was totally blind (which was recorded in the data collection form). The same tests were conducted in Health 2000 Survey (Rudanko and Koskinen 2008).

For the examination of near vision, the subjects held the chart at a distance where they could see it best. The subjects were asked to indicate the last line that they could still easily read. Testing was started on the line above by asking the subjects to read the letters on that line. If the subjects correctly identified all those letters or at least four letters of five, they were asked to move one line down towards smaller letters. The result was the lowest line on which the subject correctly identified at least four letters. If the subject was unable to see even the biggest letters, the result was entered as 99.

For the examination of distant vision, the subject sat in a chair at four meters' distance from the chart, with eyes at the level of the chart. As in the near vision test, the result was entered as the lowest line on which the subject correctly identified at least four letters. If the subject was unable to see even the biggest letters, the result was entered as 99.

Subjects with visual acuity values of 0.40 or less for near vision or 0.80 or less for distant vision were asked further questions about whether and where they had previously had their eyesight examined. If their eyesight had never been examined before, the subjects were urged to contact an optician or eye specialist. If distant vision acuity was 0.25 or lower, the subject was asked about rehabilitation services for the visually impaired.

14.3 Psychological functioning

Anna-Mari Aalto and Marko Elovainio

The concept of psychological functioning covers the individual's capacity of perceiving, interpreting and controlling him/herself and his/her environment, planning and making choices

and decisions regarding his/her life. Psychological functioning (see also Chapter 16) can be operationalized in terms of mood, personality or psychological coping resources. In the Health 2011 Survey personality dimension of psychological functioning was measured using Sense of Coherence (Antonovsky 1985) and Cynical Hostility scales (Greenglass and Julkunen 1989, Greenglass and Julkunen 1991), which are described in more detail below. The mood dimension was measured by GHQ-12 (Goldberg 1972) and BDI-13 (Beck and Beck 1972) which are described in Chapter 12.

Table 14.3.1. The position of methods used to assess psychological functioning.

Indicator	Questionnaire 1	Questionnaire 2	Questionnaire 3
Sence of Coherence (SOC)		Kys3_K1701- T11_Kys2_K55_13	
Cynical hostility			Kys1_K8101- Kys1_K8108
Beck Depression Inventory (BDI-13)[1)]		Kys1_K82S101- Kys1_K82S1804	
General Health Questionnaire (GHQ-12)[1)]	Kys1_K69- Kys1_K80		
Social support[2)]	Kys1_K680101- Kys1_K680407		

[1)] Described in Chapter 12
[2)] Described in Chapter 16

Sense of coherence (SOC) is a concept based on Antonowsky's Salutogenic Model of Health (Antonovsky 1985) and it refers to the extent to which an individual perceives the world as comprehensible, manageable and meaningful. According to Antonowsky (1985) SOC is relatively dispositional characteristic of an individual, which reflects his/her capacity to respond to challenging situations and resolving them. SOC has been associated with self-rated health and particularly with mental health (Eriksson and Lindström 2006). Antonowsky's original SOC-scale included 29-items (Antonovsky 1987), but in the present study SOC was measured by a 13-item version which was included in Questionnaire 2 (Eriksson and Lindström 2005). The participants responded on a 7-point scale to the items describing (1)

meaningfulness (e.g. *"Until now your life has had: 1=no clear goals or purpose at all ... 7=very clear goals or purpose*), (2) comprehensibility (e.g. *"Do you have the feeling that you are in an unfamiliar situation and don't know what to do? 1=very often ... 7=very seldom or never*), and (3) manageability (e.g. *"How often do you have feelings that you're not sure you can keep under control? 1=very often .... 7=very seldom or never*). Five items were reverse coded (items 2, 3, 4, 7 and 9). A total composite SOC score ranged from 13 (weak SOC) to 91 (strong SOC).

Cynical hostility is a multidimensional individual disposition comprising of affective, cognitive and behavioural components. Cynical hostility has been associated with various health problems such as somatic and psychological symptoms (Christensen et al. 2004), heart diseases risks and mortality (Nelson et al. 2004, Tindle et al. 2009, Haukkala et al. 2010) and risk behaviours (Pulkki et al. 2003). Cynical hostility was measured in Questionnaire 3 using an 8-item scale modified by Greenglass and Julkunen (1989) from the original Cook-Medley Hostility Scale (Cook and Medley 1954). The participants responded on a 4-point scale (1=fully correct, 2=quite correct, 3=quite incorrect, 4=fully incorrect) to items such as *"Most people are ready to use any means, also dishonest ones, in order to gain benefits*". A total composite sum score was calculated ranging from 8 (low hostility) to 32 (high hostility).

Appendix 5 shows the questions of psychological functioning that were presented to the young adults.

14.4 Cognitive functioning

Annamari Tuulio-Henriksson and Päivi Sainio

In the health examination, cognitive functioning was assessed with selected tasks from the CERAD neuropsychological test battery, originally developed for screening early phases of dementia and memory disturbances (Morris et al. 1989, Hänninen et al. 1999, Pulliainen et al. 1999). The cognitive functions assessed were verbal fluency, and encoding and retaining verbal material. As a part of the interview, a short version of the Mini-Mental State Examination (Folstein et al. 1975) was administered, after which a few questions concerning self-evaluation of memory function, concentration and learning new things were presented (Table 14.4.1).

Table 14.4.1. The position of methods used to assess cognitive functioning.

Indicator	Interview	Health examination	COURAGE[1)]
MMSE	HF01_K-HF09_K_YHD		
Self-evaluation of memory, concentration and learning new things	HF09_A-HF14		x
CERAD		tests on verbal fluency, word list memory and recall	
Digit span – digits forward and backward, verbal working memory			x

MMSE = the Mini-Mental State Examination; CERAD = The Consortium to Establish a Registry for Alzheimer's Disease
[1)] Described in Chapter 19.2

Cognitive tests

The tests were not performed if the subject's mother tongue was other than Finnish or Swedish, or if the subject had severe cognitive dysfunction that hindered the testing (such as severe dementia or mental disorder). The reason for not performing the test was recorded in the data collection program. The same tests were conducted in the Health 2000 Survey as well (Suutama et al. 2008).

Mini-Mental State Examination

The interview included a short version of the Mini-Mental State Examination (MMSE, Folstein et al. 1975), providing a rough overall estimate of cognitive functioning (see Appendix 4: section HF). MMSE is used in clinical practice for the detection of dementia. MMSE was administered to subjects aged 55 years or over. The test included questions on orientation and verbal memory, a mental arithmetic task and a drawing task to measure cognitive perception.

Verbal fluency

In the test of verbal fluency, the subject was asked to say aloud as many animals as possible in one minute. The study nurse measured the time with a stopwatch and kept a tally to count the number of correctly and incorrectly cited animals, as well as any repetitions of the same animal. The number of correctly and incorrectly cited animals was recorded separately in the data collection program.

Word list memory and word list recall

The subjects were shown 10 words one after another that they were to read aloud and commit to memory. After this the subjects were asked to say aloud the words they remembered; 90 seconds to given to recall the words. Then, they read the words twice again, in a different order. After each round the subjects said aloud the words they recalled. The number of words correctly and incorrectly recalled after each showing was recorded on a separate form and in the data collection program. If the subject was unable to read aloud the words, the study nurse read them out loud. The delayed recall of the words was tested by asking the subjects to repeat the same words about five minutes later, after the grip strength test had been conducted.

Self-evaluation of memory, concentration and learning new things

Self-evaluation of memory, functioning, concentration and learning new things was posed to persons aged 30 years or over. The respondents were asked to evaluate their ability in these functions using a 5-point response scale. They were also asked to assess whether their memory had changed and whether their memory problems impeded daily life. Subjects who reported that their memory had become worse were asked what they thought of being the reason for their memory problems. Repeatability of the question on memory was fair, kappa 0.68, see Table 4.2.3 in Chapter 4.2). For young adults, the questions concerning memory, concentration and learning were presented.

14.5 Social functioning

Päivi Sainio and Pirjo Tiikkainen

Social functioning generally concerns relationships with other persons and participation. It is strongly linked with the environment where one lives. Social functioning can be divided in the following dimensions: 1) social networks, 2) social activity and participation, 3) loneliness, 4) social support, and 5) social skills (Tiikkainen and Heikkinen 2011).

The Health 2011 Survey includes several questions that can be applied as indicators of social functioning (Table 14.5.1) although some of them can be used within other frameworks and using other concepts as well. The questions derive from different traditions to define and measure social functioning (Weiss 1973, Berkman and Syme 1979, Jylhä and Aro 1989, Bowling 1997, Heikkinen 1997, Glass et al. 1999).

Appendices 4 and 5 give the information which questions concerning social functioning were presented to the young adults (18–28 years). Furthermore, the COURAGE sub-study included a large section on social networks (see Chapter 19.2).

Table 14.5.1. Indicators of social functioning and their position in Health 2011 Survey.

Indicator	Interview	Questionnaires
1) Social networks		
The size of social networks		
Number of children	BC02 (men),	
The size of household	BD22 (women)	
Living alone	AB01	
Marital status	AB01	
	AA01	
Contacts		
Seeing friends and relatives		Q1: Kys1_K2010, Kys1_K2012
2) Social activity and participation		
Participation in leisure time activities		Q1: Kys1_K2001-Kys1_K2014[1)]
Health problems hindering leisure time activities		Q1: Kys1_K0501
Musculoskeletal symptoms hindering leisure time activities[2)]		Q2: OIRE_65B
Productive social participation		
Helping others	HD09-10	
3) Loneliness[3)]		Q1: T11_Kys1_K52
4) Social support		
Possibilities to get help and support from people close to oneself		Q1: Kys1_K680101-Kys1_K680407[4)]
Use of help	HD-section[5)]	
Satisfaction with personal relationships[6)]		Q1: T11_Kys1_K42_4
5) Social skills		
Taking care of matters together with other people	HA01N	
Presenting matters to unknown people	HA01O	

Q1 = Questionnaire 1; Q2 = Questionnaire 2
[1)] Described in Chapter 16
[2)] Repeatability: kappa 0.81 (see Table 4.2.1 in Chapter 4.2)
[3)] Repeatability: kappa 0.75 (see Table 4.2.2 in Chapter 4.2)
[4)] Described in Chapter 16
[5)] Described in Chapter 15.2
[6)] The question belongs to EUROHIS-8 quality of life -measure

14.6 Basic and usual activities of daily living

Päivi Sainio

In the interview, daily and usual activities (activities of daily living ADL, and instrumental activities of daily living IADL) were assessed with a question: *"How do you manage the following activities nowadays?"* (see Appendix 4, Section HA and Table 14.6.1). The ADL items were: getting in and out of bed, dressing and undressing, eating, bathing, and toileting based on (Katz et al. 1963, Katz et al. 1970). The IADL items were: using telephone, taking medication, shopping, preparing meals, doing laundry, heavy cleaning, carrying shopping bag, managing money and handling matters in public offices based on (Lawton and Brody 1969). The response categories were: no difficulties, some difficulties, much difficulties, not able. Questions on ADL and IADL were slightly modified from those used in the Mini-Finland Survey (Aromaa et al. 1989) but the topics were similar. Some updating was carried out in order to make some of the questions better reflect activities in the present-day environment.

Table 14.6.1. The position of methods used to assess usual activities in the Health 2011 Survey.

Indicator	Interview	Questionnaires	COURAGE[1)]
ADL	HA01-HA01F	Q1: Kys1_K0102 Q3: Kys3_SYOM, Kys3_PUHU, Kys3_ERIT	x
IADL	HA01G-HA01F, HA02H-HA02M	Q1: Kys1_K0103, Kys1_K0502 Q2: T11_Kys2_K47 Q3: Kys3_TAVA	x

ADL = activities of daily living; IADL = instrumental activities of daily living IADL; Q1 = Questionnaire 1; Q2 = Questionnaire 2; Q3 = Questionnaire 3

[1)] Described in Chapter 19.2

In the questionnaires there were a few additional questions concerning the daily and usual activities. The EQ-5D quality of life instrument (see Chapter 17) includes questions about managing self-care and usual activities, with a 3-point response scale (Questionnaire 1). The 15D measure of quality of life includes a couple questions on (I)ADL-functioning (Questionnaire 3, see Chapter 17). In Questionnaire 1, those who had a chronic illness or injury, were asked if this hindered their household chores. Answers were given with 10-point

scale with no hindrance to very significant hindrance. Questionnaire 2 included a question inquiring if back, neck, shoulder or joint pain within the past 30 days had prevented the person from performing daily tasks or duties.

Repeatability of the question on shopping was fair (kappa 0.76). The question on musculoskeletal problems preventing daily tasks also had fair repeatability (kappa 0.62, see Table 4.2.2 in Chapter 4.2).

14.7 Comparability of the results between Health 2000 and Health 2011 Surveys

Päivi Sainio, Tommi Härkänen and Sari Stenholm

In the performance tests, there was some variation between the nurses in how they recorded the performance of subjects who were unable to perform the test due to disability, e.g. unable to stand up from a chair or walk. The nurses were instructed to record the inability in the appropriate field in the data collection program, but in some occasions they left the field empty. In these cases there was a compulsory field to be filled with the reason for not conducting the test. During the data checking phase, all these notes from the performance tests conducted in different examination stations were used to complement missing information. The checking and possible corrections were made according to the rules used in correcting data of the Health 2000 Survey, in order to maintain the comparability. The same principles were followed when checking the data of cognitive tests.

Some modifications in the physical and cognitive tests may have influenced the comparability of the results between the Health 2011 and Health 2000 Surveys (Table 14.7.1). The variables directly comparable between the surveys are named equally by the name used in the Health 2000 Survey. In case of a minor modification in the test protocol the T11-ending was added to the otherwise equal variable name. Health 2011 variables which did not have a corresponding variable in the Health 2000 Survey were named differently.

Table 14.7.1. Modifications in the test protocols of the Health 2011 Survey compared to the Health 2000 Survey.

Test	Modification in the test protocol
Walk test (6,1 m)	A split time was taken at 4 meters Some differences in recording the values for those unable to walk Conducted for subjects aged 30 years or older (in Health 2000 55 years or older)
Chair stand	10 stands with a split time at 5 stands (in Health 2000 only 5 stands) Some differences in recording the values for those unable to stand up Conducted for subjects aged 30 years or older (in Health 2000 55 years or older)
Joint functions	Only abduction of the arms and squatting were conducted In the abduction of the arms, each arm was rated separately
Balance (Guralnik's test)	The test was conducted only for subjects aged 70 years or older The test protocol was different.
Hand grip strength	Different dynamometer was used. See below the analysis of the agreement.
Cognitive tests	Differences in recording to whom the test was not conducted (e.g. those not Finnish or Swedish speaking or those with severe dementia) The word list was shown three times for all subjects (in Health 2000 the test was ended if the subject learned all words at first or second round)

Hand grip strength – comparability with the Health 2000 Survey results

Grip strength in the Health 2011 Survey was measured using a different dynamometer (Jamar/Saehan) than in the Health 2000 Survey (Good Strength, Metitur Oy) because Metitur's dynamometers were no longer available. To assess the comparability of the two devices, 40 volunteers (20 women and 20 men) were measured with both devices. The statistical comparison was conducted using the Ordinary least products (OLP) regression method (Ludbrook 2002), in which the Jamar device was used as the independent and Metitur as the dependent variable. Confidence intervals were calculated using the bootstrap method. This OLP method gives separate estimates for fixed bias and proportional bias thus avoiding problems related to intraclass correlations (ICC).

As the confidence intervals corresponding to fixed bias (the intercept terms) contain zero and all confidence intervals corresponding to proportional bias (the slope terms) contain one (Table

14.7.2), there is no disagreement between the two devices. Fixed bias is negative for the full data indicating lower values of Metitur measurements for small values of Jamar/Saehan measurements, and for women the estimate for proportional bias is below one (0.86) indicating that for larger values of Jamar/Saehan measurements the Metitur measurements are smaller (Figures 14.7.1 and 14.7.2), but these differences are not statistically significant. The ICC estimates vary from 0.81 (only women) to 0.96 (men and women), indicating fair to excellent agreement.

Table 14.7.2. Comparison between two hand grip devices among 40 volunteers.

Group	Bias	Estimate	2.5%	97.5%
All	Fixed	-37.64	-83.00	2.17
All	Proportional	1.07	0.97	1.18
Men	Fixed	8.18	-157.77	121.76
Men	Proportional	0.99	0.77	1.31
Women	Fixed	27.05	-42.55	92.32
Women	Proportional	0.86	0.67	1.06

Figure 14.7.1. Scatter plot for the observed measurements (boxes). Solid lines correspond to fitted regression lines for full data (black), men (blue) or women (red). The dashed line corresponds to the diagonal (if the devices gave identical results, all points would lie on this line).

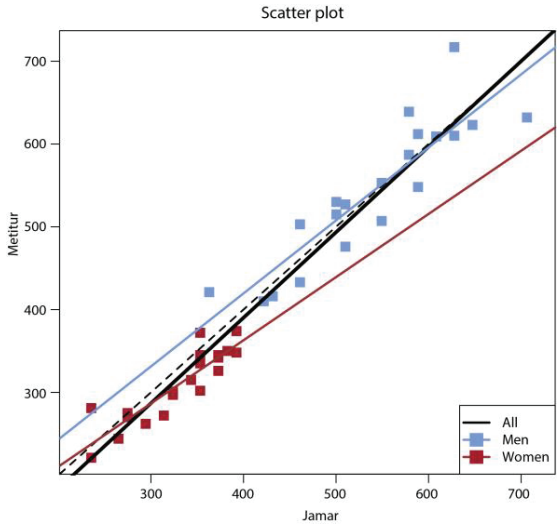
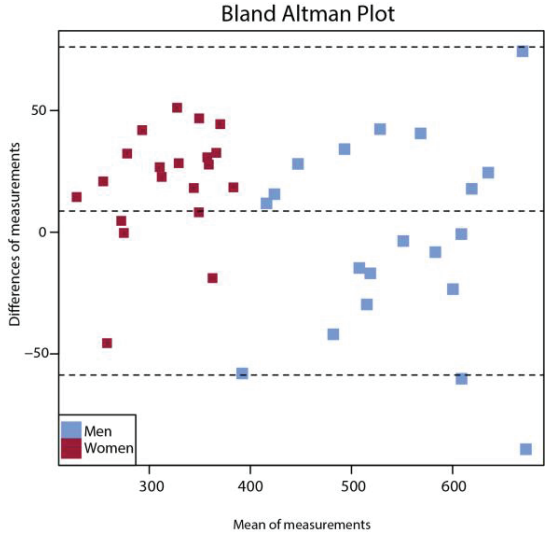


Figure 14.7.2. The Bland Altman plot for the hand grip measurements (in N). Red squares correspond to women and blue squares to men.



15. USE OF ASSISTIVE DEVICES AND NEED AND USE OF ASSISTANCE

Päivi Sainio

15.1 Use of assistive devices

The use of assistive devices was assessed in the interview (see Appendix 4). The questions concerned assistive devices for vision, hearing, communicating, mobility, special aids for eating or other daily activities, and supports, braces or prostheses. The respondents were also asked if they needed any other assistive devices or aids due to disability.

Questions on assistive devices for vision, hearing and mobility were included in the phone interview as well as the short questionnaire (see Appendix 4). Questionnaire for young adults only included a question on use of glasses (see Appendices 4 and 5).

15.2 Need and use of assistance and helping others

In the interview, the need and use of repeated assistance or help in everyday activities (for example household work, washing up, shopping etc.) due to disability was assessed (see Appendix 4). Those receiving help were further asked in which activities they require help, from whom they had received it, and how often. They were also asked if they had an informal caregiver (receiving allowance) and if they used services like meal services, transport services or bathing or sauna service outside home. In addition the participants were inquired if they had been in interval care in a nursing home or hospital ward during the past 12 months. Finally, they were asked to assess if they received enough help to manage at home and if no, in which activities would they need help and how often.

Those not receiving help in everyday activities were asked if they needed such help, and if yes, in which activities help would be needed and how often.

The respondents were also inquired if they provided regular help for other people (relatives, friends, neighbours) in tasks that help them to manage at home. Those helping others were asked who they helped, in which activities, and how often. In addition, they were asked about if they lived together with the person they helped and if they were this person's informal caregiver receiving allowance.

An abbreviated set of questions regarding need and use of help and helping others were included in the phone interview and short questionnaire (see Appendices 4 and 5). The questions were not included in the questionnaire for young adults.

16. SOCIAL CAPITAL AND PARTICIPATION

Tarja Nieminen and Tuija Martelin

The concept of social capital has been defined in many different ways but most of the definitions include the same elements. In general it refers to social structure or social networks which are characterized by norms of trust and reciprocity. Social capital has been suggested to facilitate coordination and co-operation in society and to increase well-being and health. For example, there is evidence of an association between social capital and self-rated health, cardiovascular diseases, mental health, and mortality. There are two major approaches to the level of defining social capital: some scholars see it as a collective property, while others see it as an individual property. Regardless of the approach, social capital is considered a multidimensional phenomenon. The core dimensions include social networks, social participation, social trust and reciprocity. Earlier, also social support was often seen as one dimension but in the recent literature social support has mostly been excluded from the analyses, particularly if social capital is understood as broader networks which provide their members with new information and opportunities (Nieminen 2015).

The self-administered questionnaires in Health 2011 Survey include several questions that can be applied as indicators of social capital, although they can be used within other frameworks as well (e.g. social and psychological functioning, see Chapter 14). Social networks and participation are mapped in Questionnaire 1 by means of a question concerning leisure time activities: *"How often do you practice the following activities on an average?"* (response options: every day or during most days, once or twice a week, once or twice a month, once or a few times a year, less frequently or never). This battery of questions includes nine items, all more or less relevant from the point of view of social capital: club or society activities; theatre, movies etc.; studying; church or other religious activities; exercise, hunting, fishing or other outdoor activities; handicrafts, playing music, singing etc.; visiting family, friends or neighbors; having family, friends or neighbors visit you; talking on the phone.

The cynical mistrust scale, included in Questionnaire 3, can be used to indicate *social trust and reciprocity* (see also Chapter 14.3). The scale is a short version of the Cook–Medley hostility scale (Cook and Medley 1954, Greenglass and Julkunen 1989) and it consists of eight items. The respondent is asked to read eight claims and circle the option closest to the truth (response

options: fully correct, quite correct, quite incorrect, and fully incorrect). In addition to applying the sum score based on the entire scale, two items can be considered particularly relevant from the point of view of social capital: 'It is better not to trust anyone' (social trust) and 'Most people would not want to go through the trouble to help other people' (reciprocity). Moreover, questions concerning the perceived safety of the neighbourhood and feeling safe walking out alone after 10 PM (Questionnaire 1, see Chapter 8.4) can also be interpreted to indicate trust.

Social support was measured in Questionnaire 1 by means of a set of four questions, based on the more extensive Social Support Questionnaire developed by Sarason et al. (1983, see also Chapter 14.5). The respondents were asked to estimate their possibilities to get help from people close to them when in need of help or support. Several sources of help were listed (husband, wife or partner; other relative; friend etc.), and the respondent could choose several options. The four items were: *1) "On whose help can you really count when you feel exhausted and need relaxation?"; 2) "Who do you think really cares about you no matter what happened to you?"; 3) "Who can really make you feel better when you feel down?"; and 4) "From whom do you get practical help when needed?"*. In the analyses these questions have been used in different ways, e.g. by first categorizing them according to the number of sources of support and then constructing a summary measure by means of factor analysis (see e.g. Nieminen et al. 2008).

17. QUALITY OF LIFE

Samuli Saarni and Annamari Lundqvist

Quality of life (QOL) refers to a broad, multidimensional concept that usually includes subjective evaluations of both positive and negative aspects of life while health related quality of life (HRQOL) aims to capture the aspects of QOL that can be influenced by health and health care. These include domains related to physical, mental, emotional, and social functioning. Several methods to assess the QOL and the HRQOL exist.

In the Health 2011 Survey, the EUROHIS-QOL 8-item index (Power 2003) was included in Questionnaire 1. The EUROHIS-QOL 8-item index is composed of eight items (overall QOL, general health, energy, daily life activities, esteem, relationships, finances, and home) taken from the WHOQOL-BREF. The index has the same response scale as the WHOQOL-BREF; that is, each question has an individualized five-point scale. Each item is scored positively. The overall QOL score is formed by a simple summation of scores on the eight items, with higher scores indicating better QOL.

The HRQOL was assessed by EQ-5D measure and the 15D measure, both included in Questionnaire 3.

The 15D includes 15 questions, tapping 15 dimensions of HRQOL: mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort and symptoms, depression, distress, vitality and sexual activity (Sintonen 1994, Sintonen 1995). Each dimension has five grades of severity, and so the 15D defines a vast number of health states. The 15D can be used as a single index score measure or as a profile. To calculate the 15D utility index, valuations were elicited from the Finnish population using the multi-attribute utility method. Subjects with 12 or more completed 15D dimensions were included, and missing values were predicted with linear regression analysis using the other 15D dimensions, with age and sex as independent variables, as recommended by the author of the measure (Sintonen 1994).

The EQ-5D measure (Rabin and de Charro 2001) which is among the most evaluated HRQOL measures includes five questions, tapping five dimensions of HRQOL: mobility, self-care, usual activities, pain or discomfort and anxiety or depression. Each dimension is divided into three categories of severity: no, moderate or extreme problems. Thus the EQ-5D defines 243 different health states. It too can be converted into a single score representing health utilities. Only respondents fully completing the instrument were included in the analysis; there is no established way to reliably impute missing EQ-5D values.

See TOIMIA-database (http://www.thl.fi/toimia/tietokanta/suositus/40/) for more information on the usability, validity and reliability of the quality of life –measures used in Health 2011 Survey.

18. HEALTH SERVICES AND HEALTH PROMOTION

Päivikki Koponen

18.1 Need, use and experiences of health services

In the Health 2011 Survey most questions on health services and health promotion were identical to or modified versions of questions included in previous Finnish surveys. Most of them had also been used in the Health 2000 Survey. Some new questions on patient experiences were added and some questions were dropped or edited due to minor changes in the Finnish health service system. Questions on health services were presented in the interview. Key questions on self-perceived need for care, outpatient visits, oral and mental health care as well as health examinations were also included in the phone interview and the short questionnaire as well as in the self-administered questionnaire for the young adults.

After the questions on diseases (see Chapter 11) in the interview, participants were asked if they needed continuous medical care, and if they had not received such care, which factors had hampered their access to such care (e.g. long waiting times, high costs).

Questions on health service utilization concerned in-patient hospital care within the last 12 months and the respondents' primary doctor and primary nurse. Next, they were asked about ambulatory visits to a doctor due to illnesses during the past 12 months. A separate item was included on phone and internet contacts. Corresponding questions were presented on contacts with a public health nurse. If the respondents had visited a doctor within the past 12 months, they were asked about the reason and place (e.g. public or private clinics) for their latest visit as well as their experiences on this latest visit (e.g. getting enough information and help).

A series of questions was included on the use of mental health services during the past 12 months (e.g. in which organisation, public or private, they had visited and which professional, medical doctor, nurse or psychologist, they had seen, and which kind of care, therapy or

medication, they had received). They were also asked about their experiences of the mental health care they had received (in the same way as for visits to a doctor).

The interviewees were also asked about their participation in various health examinations during the past five years; examples include the examination needed for the driving license medical certificate, health examination in occupational or student health care, health examination for war veterans, and examinations related to pregnancy or contraceptive use. The respondents were then asked about specific screening examinations such as those for eyesight, hearing, cholesterol, blood glucose, bone density, colorectal cancer screening and HIV test. Women were asked about mammography and cervical cancer screening (pap test) and men about prostate cancer screening. The screening participation was not specified to organized screening, i.e. including sporadic tests. Those aged 70 years or older were asked about home visits to assess their need for services.

Questions on physical therapy inquired the number of visits over the past 12 months, the provider of the therapy as well as experiences of the therapy, if they had received physical therapy within the past 12 months. Questions on other therapies included massage, occupational and speech therapy and alternative therapies such as chiropractic, homeopathic or lymph therapy/treatments, acupuncture and homeopathy.

Questions on oral health care were asked after questions on oral health, including visits to dentists and dental hygienists, dental and oral examinations and treatments, as well as experiences on the latest visit to a dentist.

Use of rehabilitation services was asked after questions on work and work ability, including questions on having attended occupational rehabilitation or other rehabilitation aiming to reduce limitations in functional abilities within the past 10 years (or 12 months), and experienced need for such rehabilitation.

18.2 Use of medicines

Since the 1960s (Heinonen 1966, Purola et al. 1967) respondents have been asked in Finnish national health surveys to indicate the names of their medicines, and the responses have been checked by the interviewer from prescriptions or packages. The same method was used in the Health 2011 Survey. The respondents were asked if they had used prescription drugs during the

past 12 months, and if they were currently using prescribed medicines, the names of those medicines were recorded, and their use during the past seven days was specified. Next, the use of non-prescription drugs, including vitamins, natural medicines, and homeopathic products during the last 12 months was asked and the names of such products were recorded when they were used currently. Medicines were classified in accordance with the ATC classification system (ATC DDD 2011), other products were coded using a code K (an indefinite medicine, Finnish Statistics on Medicines 2010)). Only prescribed medicines were asked in the postal questionnaire for young adults.

18.3 Health promotion

Participation in health promotion group activities and courses was assessed with a question in the self-administered Questionnaire 1 (see Appendix 5) mailed to all who were invited to the health examinations. This question was adapted from the question used in the Health 2000 survey by updating the activities, including weight control, smoking cessation, fitness training, relaxation and stress relief (e.g. yoga and meditation), alcohol and other addiction control, self-care and mental health support.

19. CONTENTS OF SUB-STUDIES CONDUCTED IN THE HEALTH 2011 SURVEY

The Health 2011 Survey included several sub-studies. The samples for sub-studies are described in Chapter 1.2.

19.1 Physical Activity and Fitness

Tommi Vasankari, Pauliina Husu, Jaana Suni, Harri Sievänen, Heli Valkeinen, Katja Borodulin and Tomi Mäki-Opas

Global recommendations of physical activity for health (WHO 2010) suggest that adults should do at least 150 minutes of moderate-intensity physical activity or 75 minutes of vigorous-intensity physical activity throughout a week or equivalent combinations of moderate- and vigorous-intensity physical activity. Moreover, muscle-strengthening activities should be done involving major muscle groups on two or more days a week. The health and welfare benefits of physical fitness are well acknowledged (Kodama et al. 2009). The role of sedentary behaviours on health and welfare as well as on societal costs has also been recently discussed (Kolu et al. 2014, Borodulin et al. 2015a, Sosiaali- ja terveysministeriö 2015). The current development in measuring physical activity and sedentary behaviours emphasizes the role of objective measurements of physical activity, obtained by accelerometers, for example. However, regardless of technological development, there is still a need for supplementing the information collected by objective measurements using questionnaires and physical activity diaries.

As part of the Health 2011 Survey, a sub-sample of individuals was invited to the Physical Activity and Fitness sub-study. All participants went through physical fitness tests at the end of the health examination and used an accelerometer during seven consecutive days beginning from the health examination. The Physical activity and fitness study was conducted in collaboration between UKK Institute and THL.

Sport and exercise questionnaire

Participants in the Physical Activity and Fitness sub-study were asked to fill in a physical activity questionnaire (form T4017) that included questions on environment and mobility, work and study environment, walking and cycling, social support, and regular sport and exercise behaviours.

Objectively measured physical activity

Participants were instructed to use a tri-axial accelerometer (Hookie AM 20, Hookie Technologies Ltd, Espoo, Finland) during the seven consecutive days beginning from the health examination. The accelerometer was attached to a waist-mounted elastic belt and placed on the right side of the hip before fitness testing at the health examination site. The participants were advised to use the accelerometer together with the belt during the waking hours excluding swimming, having a shower and bathing. The accelerometer was not used during sleeping. Written information on how to use the accelerometer was also given. After seven days of use, the accelerometer and diary (see below) were mailed to the UKK Institute where the recorded accelerometer data was analysed.

Physical activity diary

The participants received a physical activity diary (form T4018), where they were asked to report their daily physical activities, exercise, and commuting physical activity. Participants were asked to report the type of exercise or physical activity as well as the start and end times of each activity at one minute accuracy. Participants were also asked to estimate the perceived strain of exercise and physical activity using a scale: 1) light activity (no sweating or breathlessness); 2) moderate activity (some sweating or breathlessness) and 3) vigorous activity (with strong sweating or breathlessness).

Physical fitness tests

In the invitation letter the participants were instructed to wear light indoor sportswear for the fitness tests. A heart rate monitor (Polar Electro Ltd, Kempele, Finland) was used during the tests. The chest strap of the monitor was fastened around the participant's chest and the wristband was attached to the right wrist.

Before the physical fitness tests the participants were asked to fill in a short screening questionnaire (form T4318) including questions on physical activity, cardiorespiratory and musculoskeletal symptoms and diseases as well as chest pain and dizziness during exercise. The aim of the questionnaire was to identify the subjects for whom physical fitness tests might be contraindicated.

The order of physical fitness tests was:

1) One leg stand test (Suni et al. 1996),

2) Jump and reach (vertical jump) test (Suni et al. 1996),

3) Modified push-up test (Suni et al. 1996) or dynamic trunk flexion test (Oja and Tuxworth 1995), and

4) Six-minute walk test (American Thoracic Society 2002).

One leg stand test

One leg stand test measures static postural control while the area of support is reduced (Suni et al. 1996). The contraindications to the test were severe dizziness, low back pain and lower limb symptoms that might be aggravated by the test. The test was done using sport shoes or other comfortable low heel shoes. The participant chose the preferred leg to stand on. The heel of the opposite foot was placed against the inner side of the supporting leg (heel at the level of the knee joint) so that the thigh was rotated outwards. Both arms were hanging relaxed at the side of the participant's body. The participants were asked to keep their eyes open. The maximal duration of the test was 60 seconds. A nurse started timing when the participant had reached the correct test position. Timing stopped when the participant lost balance (i.e. foot of the free leg lost contact with the supporting leg or the supporting leg moved) or when 60 seconds was reached. Two attempts were performed unless the result of the first trial was 60 seconds. The recorded test result was the longest time in seconds the participant maintained the correct test position (0–60 s).

Jump and reach test

Jump and reach test was used to measure maximal power of the lower limbs. Contraindications included severe obesity, severe dizziness or severe symptoms of the spine or the extremities that might be aggravated by the test movements. The aim was to jump as high as possible. The participant stood beside the jump-board facing forward. The dominant upper extremity was raised up straight against the jumping board. The standing height was marked with the magnesium powdered middle finger. After that, vertical jumps were performed. One practice trial and two test trials with maximal effort were performed. The vertical difference between the standing height and the jumping height was measured in centimetres with a tape measure. The highest jump height in centimetres was the recorded score.

Modified push-ups

The modified push-ups test was used to measure short-term endurance capacity of the upper extremity extensor muscles and the ability to stabilize the trunk. Contraindications included a moderate to severe disease or symptoms of the circulatory system, and severe symptoms of the lumbar region or the extremities that might be aggravated by the test movements. The participants lied prone on the mat, and began the push-up cycle by clapping hands behind the back once; this was followed by a normal straight-leg push-up with elbows completely straight in the up-position, so that the participants could touch their either hand with the other hand. The participants ended the cycle in prone position. Position of the arms was controlled. The different phases of the modified push-up were practiced once before the test trial. There was one test trial. The number of correctly performed push-ups completed in 40 seconds was counted. The recorded test score was the total number of correctly performed push-ups.

Dynamic sit-ups test

The dynamic sit-ups test was used as a secondary test for those participants who were unable to perform the modified push-ups test. The test measured the dynamic strength of the abdominal and hip flexor muscles. Contraindications included severe pain of spine or hips that might be aggravated by the test movements. The participant laid supine on the gym mat, knees flexed to 90 degrees (knees and ankles together). The tester supported ankles with his/her hands so that during the test performances the feet of the participant stayed on the gym mat.

The test was performed in four performance levels with increasing difficulty. The aim was to perform 5 sit-ups in each level without a rest between the levels. The movement should be smooth and the back of the head and elbows should touch the mat between each sit-up:

1) flex the upper body just enough to arise the shoulder blades from the mat

2) reach mid-patella with fingertips of both hands from a straight lying position while keeping the arms straight and palms resting on thighs

3) reach thighs with both elbows while arms are folded over chest

4) reach thighs with elbows while the fingertips touch the back of earlobes.

There was one test attempt in each performance level. The test score was the number of correctly performed sit-ups (0-20).

The six-minute walk test

The six-minute walk test was used to measure cardiorespiratory fitness submaximally (American Thoracic Society 2002). Contraindications to the test were chest pain during physical activity, myocardial infarction during the previous month, elevated heart rate (>120 beats/min) or blood pressure (180/100 mmHg) at rest. The test was done on a 15 m track (typically the corridor outside the examination rooms) using cones at both ends of the track as turning points. Running and jogging was not allowed. The observer showed a fast way to make the turn and the participant walked once through the track. The observer started timing, when the participant began walking and counted the number of completed tracks during the test time. The observer asked the participant's heart rate after two and four minutes and checked it from the heart rate monitor at the end of the test. The tester told loudly when the walk had lasted 1, 2, 3, 4, and 5 minutes. The participant was informed when the last 15 seconds of the test begun and the walking was stopped exactly when the six-minute test time was over. The walking distance was calculated by multiplying the number of rounds (30 m track) plus the walked meters of the unfinished last round. The recorded result of the test was the walked distance (meters).

From objective accelerometer and physical fitness data, several summary variables have been created for researchers. The data are available for the researcher after the research proposal has been accepted by the physical activity expert group including members from THL and UKK institute. Moreover, researchers interested in utilizing objective measurements of physical activity and fitness, are encouraged to consult the physical activity expert group for more detailed measures.

19.2 Collaborative research on aging in Europe (COURAGE)

Satu Meriläinen-Porras, Ulla Laitinen, Seppo Koskinen and Päivi Sainio

The COURAGE in Europe Project collected data on the determinants of health and disability in the ageing population, with specific tools for the evaluation of the role of the built environment and social networks on health, disability, quality of life and well-being (Leonardi et al. 2014). The survey was conducted in three countries, Finland, Poland and Spain. In Finland, the COURAGE in Europe data collection was conducted in connection with the Health 2011 Survey, on a sub-sample described in Chapter 1.2.

Four main objectives have been pursued by COURAGE in Europe. The first objective was to develop valid assessment instruments to measure key health and health-related outcomes in the general population (from age 18 years to end of life). The second objective was to validate the assessment instruments to create a scientific evidence base for health and disability determinants in ageing. The third was to produce substantial innovation in ageing survey methodology, by e.g. including new instruments measuring the built environment and social networks. Finally, the COURAGE in Europe project aimed to provide cross-population analysis and a baseline for longitudinal data collection.

Implementation of the COURAGE sub-study

The COURAGE protocol included an interview and several measurements taken either at the participant's home or at the health examination.

In Finland, the COURAGE interviews were carried out at the respondents' homes about one week before the Health 2011 health examination. After the respondents had received the letter of invitation to participate to the Health 2011 Survey, the COURAGE nurses made the interview appointments with respondents by phone. If the subjects could not be contacted by phone the nurses also visited the respondents' homes without making appointments beforehand, but this option was rarely pursued because it did not prove to be an efficient way to gain more interviews. The home interview lasted approximately 2–2.5 hours and consisted of an interview and a health examination (see below). The COURAGE nurses used their own cars to travel from one respondent's home to another, following a weekly fieldwork and travel plan. In a sparsely populated country like Finland, driving distances are long and the

actual distance driven by the home-visit nurses exceeded 100 000 kilometres. All the equipment and materials needed for the COURAGE survey were transported by the nurses themselves. The COURAGE nurses' health examination package included a map, a computer, a phone, a grip strength measurement device, a sphygmomanometer, a vision test table, a measuring tape, a COURAGE weekly fieldwork and travel plan, an information sheet for the participant, two informed consent forms, two consent forms on behalf of the examinee, questionnaires in Finnish, Swedish and English and sheets of sticker labels with the code number for the paper forms. COURAGE nurses arranged their own accommodation.

All 14 COURAGE nurses were part of the five Health 2011 field teams, although their work was very independent. The COURAGE project had also one head nurse and a coordinator who were in daily contact with the nurses by phone and email.

Training of the COURAGE nurses

The training of the 14 COURAGE nurses was organized by THL and held between 25 and 29 July 2011, in collaboration with training of the Health 2011 Survey. The general training of the COURAGE nurses was conducted together with the Health 2011 fieldwork personnel. The specific training included interview and measurement training of the COURAGE protocol as well as preparative actions such as making appointments and motivating respondents. Special attention was paid to the techniques of measuring blood pressure and hand grip strength as well as of performing tests of vision and digit span. During the training, questions and ambiguities were also clarified and common courses of actions were agreed. During the fieldwork, one meeting was organized for the nurses with the coordinator and the head nurse to discuss various experiences and practical problems with.

The COURAGE pilot

A pre-pilot for the COURAGE study was organized in spring 2010 to evaluate the Built Environment and Social Network sections of the COURAGE protocol. In Finland, this was carried out by means of a discussion group consisting of seven experts from THL. In October 2010, a pilot study focusing on the protocol and length of the interview was held. A total of 70 voluntary participants with different sex, age, and social background were recruited from THL personnel as well as from the residents and clients of Riistavuori service centre in Helsinki. Based on the experiences from the pre-pilot and pilot studies in Finland, Poland and Spain, some changes to the protocol was made.

Contents of the COURAGE interview and home health examination

The protocol consisted of an interview and a health examination at the respondent's home. The interview focused on the participant´s background information, health, functioning, quality of life, use of health services and well-being as well as on the factors behind these phenomena. A special emphasis was on the built environment which was studied using self-reports (Raggi et al. 2014) and direct observations (Quintas et al. 2014): the COURAGE nurses were responsible for observing and recording facts from the subjects' outdoor environment. The built environment was audited in 10 per cent of the participants. Another emphasis in the COURAGE protocol was on tools assessing social networks (Zawisza et al. 2014). The contents of the COURAGE in Europe protocol are described in Appendix 6.

The COURAGE health examination at home consisted of the following measurements and performance tests:

- blood pressure and heart rate
- height and weight
- waist circumference
- timed walk (4 m)
- near and distant vision
- grip strength
- immediate and delayed verbal recall
- digit span forward and backward
- verbal fluency

The measurements and tests similar to Health 2011 protocol were conducted at the health examination of the Health 2011 Survey and they are described in Chapter 14. Those measurements and performance tests which deviated from the Health 2011 protocol (blood pressure, distance and near vision, digits forward and backward and grip strength) were carried out in conjunction of the COURAGE interview. These methods are described below, based on COURAGE Survey Manual (COURAGE Project Consortium 2011).

Blood pressure and heart rate

Blood pressure and heart rate were measured from the right arm whenever possible, with an automatic measuring device (Omron iC-10). The device measured blood pressure three times with 60 seconds pause in between. The systolic and diastolic blood pressure and heart rate from all measurements were recorded in the electronic data collection program. Blood pressure was

also measured in the Health 2011 Survey with a standard mercury manometer (see Chapter 7.1).

Vision tests

Both distance and near vision were tested using standard vision charts, which are useful to test the approved WHO thresholds for distance and near VA (Visual Impairment <6/18; ≥3/60; Blindness <3/60). The E chart has the capital letter "E" facing in different directions and the person being tested had to determine which direction the "E" is pointing: up, down, left, or right. The vision of each eye was recorded separately, with the other eye covered.

If the subjects used normally eyeglasses they were tested while wearing them; if they had eyeglasses but did not usually use them, the test was be performed without wearing them. The vision charts needed to be well lit and at eye level.

To test distant vision, the chart with small E-letters was shown to the subject at three meters distance. First the left eye was tested with the right eye covered with the palm of the right hand. If the subject could correctly tell which direction at least three out of four small E's were pointing (result *"6/18 or better"* was recorded on the data collection program), the testing continued with the right eye. If not, the large E's were shown. If succeeded (three out of four large E's were seen, *"6/60"*), the testing continued with the right eye. If the subject could not see the large E's at 3 meters distance, the large E's were shown at 1.5 meters distance. If succeeded (*"3/60"*), the testing continued with the right eye. If the respondent could see less than three out of four of the largest 'E' from a distance of 1.5 meters, the result *"less than 3/60"* was marked. The right eye was tested similarly.

In the near vision test the subjects were instructed to hold the E-chart as close as they wanted. At least three out of four E's had to be correct on each line before testing the next. The test started with large E's on the left eye, the other was covered with the palm. If correctly seen, the medium size E's were shown followed with the small size letters. The line with smallest letters correctly identified was marked as the result in the data collecting program. The right eye was tested similarly.

Digit span

The digits forward and backward tests required the respondents to repeat a series of numbers. The test started with a trial run with two digits. The study nurse read out loud the digits at the rate of one per second, and the subject was to repeat them in the same order. If the subject failed, a second trial was given. If correct, the next series was read. The forward digit test started with a three digit set and continued as long as the subject could correctly repeat the numbers, up to nine digits (maximum). The backward list started with two digits, up to maximum of eight digits. The result was the maximum number of digits correctly repeated either at the first or second trial in each set. If not able to repeat even the first set, the result was set 0.

Hand grip strength

Hand grip strength was measured from the dominating hand using Smedley's dynamometer. The test protocol was similar to that of the Health 2011 Survey, described in detail in Chapter 14.1.

Using the data for research purposes

The data collected in the Finland are available for research purposes in collaboration with the Health 2011 project organisation. In order to obtain access to the data, researchers must first submit a study proposal which will be reviewed by the Health 2011 Scientific Board.

19.3 Sub-study on oral health

The clinical oral health examinations were carried out in the southern and northern field examination areas (i.e. Helsinki and Oulu) and the oral radiographic examination was performed only in Helsinki. The contents of examinations are described in Chapter 13.

19.4 The Finn-Home study

Ulla Laitinen and Antti Jula

In the Finn-Home study blood pressure was measured at home by the participants themselves with a validated automatic oscillometric device (Omron model M& HEM-7211-E(V), Omron Corp.-, Kyoto, Japan) according to the protocol of the Finn-Home (Koskinen et al. 2008). A cuff measuring 14x48 cm (bladder size 13x23 cm) was used for participants with an upper arm circumference of ≤35 cm. A long 16x65 cm cuff (bladder size 15x29 cm) was used for participants with an arm circumference exceeding 35 cm.

For the self-measurements the instructions (form T4009) with a diary (form T4008) and devices were given to the participants during the health examination. They were also individually guided how to measure blood pressure in a proper way. Participants were requested to refrain from heavy exercise and to avoid cola drinks, coffee, tea, eating, and smoking for at least one hour before the measurement, and to sit five minutes with the cuff around the non-dominant upper arm before the first measurement.

The Finn-Home study participants measured their seated blood pressure twice, approximately at a two-minute interval every morning between 6 AM and 9 AM and every evening between 6 PM and 9 PM on seven consecutive days.

19.5 DNA sub-study

In the Health 2011 Survey, those subjects who did not participate already in the Health 2000 Survey or did not have a blood sample available for genetic analyses were recruited to the DNA sub-study. Blood sampling and the sample processing is described in Chapter 5.

19.6 Migraine sub-study

Mikko Kallela and Pirkko Alha

The health interview included a question on recurrent, severe headaches during the last three months, last year, or ever. If the respondents indicated having ever had repeated attacks, they were asked to complete a separate headache questionnaire. This questionnaire (form T4019) included several questions on headache severity and impact and headache attack characteristics as well as use of medicines for treatment or prevention of headaches. The formulation of the questions allows diagnosis of migraine according to the current IHS criteria (Headache Classification Subcommittee of the International Headache Society 2004). Similar questionnaires have been found to be accurate (in comparison with clinical diagnosis made by a neurologist) in diagnosing migraine (Kallela et al. 2001).

Discussion and conclusions

Tomi Mäki-Opas, Päivikki Koponen, Seppo Koskinen, Tommi Härkänen and Annamari Lundqvist

In order to promote health and welfare as well as to develop successful health and social policies, accurate, adequate and up-to-date scientific information on health and welfare and their preconditions is required. For example information is needed on the social and health care needs of the growing aging population, and how their demanding service needs can be postponed and reduced. Information is also needed on changes in the working capacity among the working aged population, how longer and healthier working careers can be promoted, and which conditions promote employment opportunities among younger age cohorts. It is imperative to evaluate how social inequalities in health and health behaviours as well as health and social welfare costs can be diminished. These are just a few examples of questions which can be answered in future analysis of the Health 2011 Survey data.

Valid information on the above mentioned phenomena can only partly be acquired from other sources such as registers or from postal surveys. The health and social services will never cover the total population even though the Finnish health care system aims to allow access to all citizens. Some persons do not need or they do not seek services. Information from patient/client records and registers usually does not reveal the complexity of problems and needs of e.g. persons with multiple health and social problems. Differences between regions and various population groups identified in registers can be biased due to differences in professional practices, for example, not revealing true differences between regions and population groups. Service records and registers rarely include any information on health risks and protective factors, even though such information is essential for defining and evaluating needs for prevention and health promotion.

Information on self-perceived need for care and on how these needs are met can only be obtained by asking the persons themselves. It is essential to evaluate how these experiences are associated with objectively measured health status. Clinical measurements and samples obtained from health examinations provide objective data without reporting bias, which is especially important in items prone to be biased by social desirability of reporting. Health

examinations are also needed to provide data on phenomena which suffer from lack of awareness by the individuals themselves and differences in diagnostic practices.

Collecting accurate information of the health and welfare as well as their determinants among different population groups requires a large and population-based random sample of the whole target population. The Health 2011 Survey included both cross-sectional and longitudinal repeated measures study designs. Cross-sectional population-based random samples are needed to describe the current trends in health and welfare as well as their determinants in Finland. The longitudinal population-based study design provides possibility to evaluate chances both in health and welfare and in their risk factors and determinants. It also provides a possibility to forecast health and welfare changes in the population with different scenarios based on the risk factors, determinants or policy alternatives.

The Health 2011 Survey was based on a representative sample of Finnish population, who had also been invited to the Health 2000 Survey while being 18 years of age or older and who thus were at least 29 years of age during the Health 2011 Survey. A new sample of young adults was also invited to the Health 2011 Survey. We assume that participation in the Health 2000 Survey did not influence the representativeness of the Health 2011 Survey as it is unlikely that a single health examination would have motivated major changes in health behaviour and service utilization after the examination. Moreover, the original sample of the Health 2000 Survey was representative of the total population in 2011, with one exception. Mortality and emigration had affected the Health 2000 sample and the rest of the population in a similar way.

The only limitation regarding the representativeness of the sample is that it does not include persons who had immigrated to Finland after the year 2000. However, a specific study conducted in 2010–2012, the Migrant Health and Wellbeing Study (Maamu), with information comparable with that obtained in the Health 2011 Survey, provides a unique opportunity to compare the health, functioning and welfare of immigrants and the general population (Castaneda et al. 2012). Comparable data from the Health 2011 Survey participants from the six cities where the Maamu study was conducted, and in the same age range as the Maamu sample (18–64 years) have been used for the general population comparison group for the three migrant groups.

With adequate measurements, interviews and questionnaires, the Health 2011 Survey was able to provide comprehensive up-to-date information on health, functioning, work ability, mental health, oral health, musculoskeletal diseases, need for services, objective and perceived measures of welfare, and the determinants of these. The measurements, interview items and

contents of the questionnaires were developed mainly by the same expert groups that were involved in the Health 2000 Survey and, moreover, the members of these expert groups were involved in training of the fieldwork personnel. Thus the experiences from the previous study were taken into account. However, many problems were faced in the planning, preparations, fieldwork and data management, mainly due to limited time, personnel and funding.

The recruitment of the fieldwork personnel was more difficult than expected. During the year 2011, the labour market situation among nurses was quite good and even after several newspaper ads, contacts with the universities of applied sciences for nurse training, and other contacts, it was hard to recruit nurses with suitable experience for the health examinations and other demanding fieldwork tasks such as recruiting and motivating participants, and carrying out home visits independently.

In spite of the intensive training period for all fieldwork personnel, too little time was available for the nurses to practice the measurements and other tasks. Although the fieldwork started generally in a fluent way, some of the key problems especially in the IT environment could have been avoided with more testing and piloting, as well as practical training before the fieldwork.

The participation rate of the Health 2011 Survey can be regarded as acceptable, even though it was much lower than in the Health 2000 Survey. Of those who were invited to the Health 2011 Survey, 66.5 per cent participated in at least one phase of the data collection. Among the sample aged 29 or over, i.e. the Health 2000 Survey sample, the participation rate was 72.6 per cent. Moreover, 75.9 per cent of those who participated in the Health 2000 Survey also took part in the Health 2011 Survey. Among the new sample of young adults, the participation rate was only 42.0 per cent. The participation rate in the Mini-Finland resurvey was 81.3 per cent, and a total 78.6 per cent of sample had participated in all three surveys (i.e. the Mini-Finland, Health 2000, and Health 2011 Surveys).

Various means to increase the participation rate were utilized, such as use of national and local media in promoting the Health 2011 Survey, repeated attempts to contact those who did not arrive at the health examination, complementary data collection after the actual fieldwork period including a contracted health examination and home visit health examination, the phone interview and postal questionnaire. Moreover, invitations to participate in the Health 2011 Survey were also sent by SMS to those subjects who could not be reached by phone. In addition, all subjects who filled the postal questionnaire sent to the non-participants received a small gift (value of 8 euros) and also took part in a lottery (value of prize about 500 euros).

The significant reduction in the participation rate, compared to the high level reached in the Health 2000 Survey (over 90 per cent participated at least in some data collection phase), is probably largely due to some essential differences between the protocols of the Health 2000 Survey and the Health 2011 Survey. In the Health 2000 Survey all sampled persons were contacted by well trained and experienced interviewers from the Statistics Finland. They also carried out the interviews and at the end of the interview, booked the examination visit at a time most suitable for the participants. The mailed invitation letter did not succeed as well in motivating the subjects to participate in the Health 2011 Survey. Also the medical examination by a physician and a clinical dental examination for all participants motivated participation in the Health 2000 Survey. These were not available in the Health 2011 Survey due to lack of resources. Future surveys should consider personalized invitations and measurements of interest among the population.

The effects of non-participation have been adjusted for using statistical methods (Härkänen et al. 2016). The most common methods are based on weighting or multiple imputation methods. These methods were shown in the work of Härkänen et al. (2016) to provide quite accurate results when comparing the true prevalence rates of disability pension, hospitalization and reimbursement of medication in the whole sample, obtained from national registers, with the prevalence estimates based on the participants. Although these results alleviated the general concern that non-participation creates biased results, the effects of non-participation must be assessed from both the statistical and substantive points of view in all analyses using the Health 2011 Survey data – as well as in analyses based on other survey data sets.

Implementation of a large health examination study is expensive. Although funding for the Health 2011 Survey was gathered from over twenty sources, several contents of the Health 2000 Survey were left out from the Health 2011 Survey or moved from the interview mode to self-administered questionnaires to save costs. In all longitudinal analyses the effects of these changes have to be taken into account. Due to the limited resources for the Health 2011 Survey, quality assurance was not as systematic and timely as in the Health 2000 Survey and as recommended (Tolonen 2013). This caused some errors and problems in the data sets, which needed more efforts than expected after the fieldwork. During the fieldwork not all problems could be detected as soon as preferred, and sometimes actions were taken to correct the detected problems later than in the ideal case. However, quality assessment of the key items and measurements showed satisfactory results. Future surveys should include at least more frequent checking of the data during the fieldwork and more communication and feedback between the main office and the fieldwork teams.

Despite the above mentioned limitations, the Health 2011 Survey provides abundant, valuable and valid data for health monitoring and multidisciplinary public health and epidemiologic research. Register based follow-up also enhances the possibilities for scientific research.

References

AALTO, A. M., ELOVAINIO, M., KIVIMAKI, M., UUTELA, A. & PIRKOLA, S. 2012. The Beck Depression Inventory and General Health Questionnaire as measures of depression in the general population: a validation study using the Composite International Diagnostic Interview as the gold standard. *Psychiatry Res,* 197**,** 163-71.

AMERICAN THORACIC SOCIETY 2002. ATS Statement: Guidelines for the Six-Minute Walk Test. *Am J Respir Crit Care Med,* 166**,** 111-17.

ANTONOVSKY, A. 1985. The life cycle, mental health and the sense of coherence. *Isr J Psychiatry Relat Sci,* 22, 273–80.

ANTONOVSKY, A. 1987. Unraveling the mystery of health. How people manage stress and stay well. Josey-Bass. San Francisco.

AROMAA, A. 2008. Content of home-visit interview. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 39-45. http://urn.fi/URN:NBN:fi-fe201204193320.

AROMAA, A., HELIÖVAARA, M., IMPIVAARA, O., KNEKT, P., MAATELA, J., JOUKAMAA, M., KLAUKKA, T., LEHTINEN, V., MELKAS, T., MÄLKIÄ, E., NYMAN, K., PAUNIO, I., REUNANEN, A., SIEVERS, K., KALIMO, E. & KALLIO, V. 1989. Terveys, toimintakyky ja hoidontarve Suomessa. Mini-Suomi-terveystutkimuksen perustulokset. Kansaneläkelaitoksen julkaisuja AL:32. Helsinki ja Turku. http://hdl.handle.net/10138/162843

AROMAA, A. & KOSKINEN, S. (eds.) 2004. Health and functional capacity in Finland. Baseline results of the Health 2000 health examination survey. Kansanterveyslaitoksen julkaisuja B: 12/2004. http://urn.fi/URN:NBN:fi-fe201204193452.

AROMAA, A., REUNANEN, A., IMPIVAARA, O., HELIÖVAARA, M., KNEKT, P. & MAATELA, J. 1985. Mini-Suomi-tutkimuksen toteutus. Osa 2. Verenkiertoelinten ja hengityselinten sairauksien tutkimusmenetelmät. Kansaneläkelaitoksen julkaisuja ML: 49. Helsinki ja Turku. http://hdl.handle.net/10138/162384

BAGBY, R. M., PARKER, J. D. & TAYLOR, G. J. 1994. The twenty-item Toronto Alexithymia Scale--I. Item selection and cross-validation of the factor structure. *J Psychosom Res,* 38**,** 23-32.

BATES, D., MAECHLER, M. & BOLKER, B. 2012. lme4: Linear mixed-effects models using S4 classes. http://CRAN.R-project.org/package=lme4.

BECK, A. T. & BECK, R. W. 1972. Screening depressed patients in family practice. A rapid technic. *Postgrad Med,* 52**,** 81-5.

BEDOGNI, G., MALAVOLTI, M., SEVERI, S., POLI, M., MUSSI, C., FANTUZZI, A. L. & BATTISTINI, N. 2002. Accuracy of an eight-point tactile-electrode impedance method in the assessment of total body water. *Eur J Clin Nutr,* 56**,** 1143-8.

BERKMAN, L. F. & SYME, S. L. 1979. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. *Am J Epidemiol,* 109, 186-204.

BORODULIN, K., KÄRKI, A., LAATIKAINEN, T., PELTONEN, M. & LUOTO, R. 2015a. Daily Sedentary Time and Risk of Cardiovascular Disease: The National FINRISK 2002 Study. *J Phys Act Health,* 12**,** 904-8.

BORODULIN, K., VARTIAINEN, E., M, P., P, J., JUOLEVI, A., T, L., MÄNNISTÖ, S., SALOMAA, V., SUNDVALL, J. & PUSKA, P. 2015b. Forty-year trends in cardiovascular risk factors in Finland. *Eur J Public Health,* 25, 539-46.

BOSY-WESTPHAL A, SCHAUTZ B, LATER W, KEHAYIAS JJ, GALLAGHER D & MJ., M. 2013. What makes a BIA equation unique? Validity of eight-electrode multifrequency BIA to estimate body composition in a healthy adult population. *J Clin Nutr,* 67, S14-21.

BOWLING, A. 1997. Measuring social networks and social support. In: A. Bowling (Ed.) Measuring health. A review of quality of life measurement scales. Open University Press. Buckingham.

CASTANEDA, A., RASK, S., KOPONEN, P., MÖLSÄ, M. & KOSKINEN, S. 2012. Maahanmuuttajien terveys ja hyvinvointi - Tutkimus venäläis-, somalialais- ja kurditaustaisista Suomessa. THL Raportti 2012/61. Tampere. http://urn.fi/URN:ISBN:978-952-245-739-4

CHRISTENSEN, U., LUND, R., DAMSGAARD, M. T., HOLSTEIN, B. E., DITLEVSEN, S., DIDERICHSEN, F., DUE, P., IVERSEN, L. & LYNCH, J. 2004. Cynical hostility, socioeconomic position, health behaviors, and symptom load: a cross-sectional analysis in a Danish population-based study. *Psychosom Med,* 66, 572-7.

COOK, W. W. & MEDLEY, D. M. 1954. Proposed hostility and pharisaic-virtue scales for the MMPI. *J App Psychol,* 38, 414–18.

COTES, I. E. 1975. Lung function. Assessment and application in medicine. 3rd edition. Blackwell Scientific. London.

COURAGE PROJECT CONSORTIUM 2011. COURAGE Survey Manual: The Collaborative Research on AGEing in Europe (COURAGE in Europe).

CSUKA, M. & MCCARTY, D. 1985. Simple method for measurement of lower extremity muscle strength. *Am J Med,* 78, 77-81.

CURB, J. D., CERIA-ULEP, C. D., RODRIGUEZ, B. L., GROVE, J., GURALNIK, J., WILLCOX, B. J., DONLON, T. A., MASAKI, K. H. & CHEN, R. 2006. Performance-based measures of physical function for high-function populations. *J Am Geriatr Soc,* 54, 737–42.

DE BRUIN, A. Z., PICAVET, H. S. J. & NOSSIKOV, A. 1996. WHO. Health Interview Surveys. Towards international harmonization of methods and instruments. WHO Regional Publications, European Series no 58. Geneva.

DEROGATIS, L. R., LIPMAN, R. S., RICKELS, K., UHLENHUTH, E. H. & COVI, L. 1974. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behav Sci,* 19**,** 1-15.

DIGGLE, P. & KENWARD., M. G. 1994. Informative drop-out in longitudinal data analysis. *J R Stat Soc Ser C Appl Stat,* 43, 49-93.

DJERF, K., LAIHO, J., LEHTONEN, R., HÄRKÄNEN, T. & KNEKT, P. 2008. Weighting and statistical analysis. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 182-200. http://urn.fi/URN:NBN:fi-fe201204193320.

ERA, P. & SAINIO, P. 2008. Perceptual-motor speed. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 95-96. http://urn.fi/URN:NBN:fi-fe201204193320.

ERIKSSON, M. & LINDSTRÖM, B. 2005. Validity of Antonovsky's sense of coherence scale: a systematic review. *J Epidemiol Community Health,* 59, 460-6.

ERIKSSON, M. & LINDSTRÖM, B. 2006. Antonovsky's sense of coherence scale and the relation with health: a systematic review. *J Epidemiol Community Health,* 60, 376-81.

FAGT, S., ANDERSEN, L. F., ANDERSSEN, S. A., BECKER, W., BORODULIN, K., FOGELHOLM, M., GROTH, M. V., GUNNARDOTTIR, I., HELAKORPI, S., KOLLE, E., MATTHIESSEN, J., ROSENLUND-SORENSEN, M., SIMONEN, R., SVEINSSON, T., TAMMELIN, T., THORGEIRSDOTTIR, H., VALSTA, L. & TROLLE, E. 2011. Nordic Monitoring of Diet, Physical Activity and Overweight. Validation of indicators. TemaNord 2011:556. Denmark.

FIATARONE, M. A., O'NEILL, E. F., RYAN, N. D., CLEMENTS, K. M., SOLARES, G. R., NELSON, M. E., ROBERTS, S. B., KEHAYIAS, J. J., LIPSITZ, L. A. & EVANS, W. J. 1994. Exercise training and nutritional supplementation for physical frailty in very elderly people. *N Engl J Med,* 330, 1769-75.

FINNISH HYPERTENSION SOCIETY WORKING GROUP 2002. Suomen Verenpaineyhdistys ry:n asettama työryhmä. Kohonneen verenpaineen hoito. *Duodecim*, 118, 110-126 (in Finnish).

FINNISH STATISTICS ON MEDICINES 2010. Finnish Medicines Agency Fimea and Social Insurance Institution. Helsinki.

FLEISS, J. L. 1981. Statistical methods for rates and proportions (2nd edition). John Wiley & Sons. New York.

FLETCHER, C. M., ELMES, P. C., FAIRBAIR, N. A. S. & WOOD, C. H. 1959. The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *BMJ,* 2, 257-66.

FOLSTEIN, M., FOLSTEIN, S. & MCHUGH, P. 1975. 'Mini-mental State': a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res,* 12, 189–98.

FRIED, L. P., HERDMAN, S., KUHN, K., RUBIN, G. & TURANO, K. 1991. Preclinical Disability: Hypotheses About the Bottom of the Iceberg. *J Aging Health,* 3, 285-300.

GLASS, T. A., DE LEON, C. F., MAROTTOLI, R. & BERKMAN, L. 1999. Population based study of social and productive activities as predictors of survival among elderly Americans. *BMJ,* 319, 478-83.

GMEL, G. & REHM, J. 2004. Measuring alcohol consumption. *Contemporary Drug Problems,* 31, 467–540.

GOLDBERG, D. P. 1972. The Detection of Psychiatric Illness by Questionnaire. Oxford University Press. London.

GOULD, R., HÄRKÄPÄÄ, K. & KOSKINEN, S. 2015. Työkyvyn arviointi väestötutkimuksessa. TOIMIA-tietokanta. http://www.thl.fi/toimia/tietokanta/suositus/22/.

GOULD, R., ILMARINEN, J., JÄRVISALO, J. & KOSKINEN, S. 2008. Dimensions of work ability. Finnish Centre for Pensions, The Social Insurance Institution, National

Public Health Institute, Finnish Institute of Occupational Health. Vaasa. http://www.etk.fi/julkaisu/dimensions-of-work-ability/

GRAUBARD, B. I. & KORN, E. L. 1999. Predictive Margins with Survey Data. *Biometrics,* 55, 652-9.

GREENGLASS, E. R. & JULKUNEN, J. 1989. Construct validity and sex differencesin Cook-Medley hostility. *Person Ind Diff,* 10, 209-18.

GREENGLASS, E. R. & JULKUNEN, J. 1991. Cook-Medley hostility, anger, and the Type A behavior pattern in Finland. *Psychol Rep,* 68, 1059-66.

GRIMBY, G., BJURE, J., AURELL, M., EKSTRÖM-JODAL, B., TIBBLIN, G. & WILHELMSEN, L. 1972. Work capacity and physilogic responses to work. *Am J Cardiol,* 30**,** 37-42.

GURALNIK, J. M., FERRUCCI, L., SIMONSICK, E. M., SALIVE, M. E. & WALLACE, R. B. 1995a. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med,* 332, 556-61.

GURALNIK, J. M., FRIED, L. P., SIMONSICK, E., M., KASPER, J. & LAFFERTY, M., E (eds). 1995b. The Women's Health and Aging Study: Health and Social Characteristics of Older Women with Disability. Bethesda, MD, National Institute on Aging, NIH Publication 1995, No. 95-4009.

GURALNIK, J. M., SIMONSICK, E. M., FERRUCCI, L., GLYNN, R. J., BERKMAN, L. F., BLAZER, D. G., SCHERR, P. A. & WALLACE, R. B. 1994. A short physical performance battery assessing lower extremity function: association with selfreported disability and prediction of mortality and nursing home admission. *J Gerontol,* 49, M85–94.

HARRELL JR, F. E. & OTHERS. 2014. Hmisc: Harrell Miscellaneous. R package version 3.14-4. http://CRAN.R-project.org/package=Hmisc

HAUKKALA, A., KONTTINEN, H., LAATIKAINEN, T., KAWACHI, I. & UUTELA, A. 2010. Hostility, Anger Control, and Anger Expression as Predictors of Cardiovascular Disease. *Psychosom Med,* 72, 556-62.

HEADACHE CLASSIFICATION SUBCOMMITTEE OF THE INTERNATIONAL HEADACHE SOCIETY 2004. The International Classification of Headache Disorders: 2nd edition. 24 Suppl**,** 9-160.

HEIKKILÄ, J. (ed.) 1982. EKG, perusteet ja tulkinta. Lääketehdas Orion. Espoo. (In Finnish).

HEIKKINEN, E. 1997. Background, design and methods of the Evergreen Project. *Journal of Aging and Physical Activity,* 6, 106-120.

HEIKKINEN, E. 1997. Background, design and methods of the project. In: Heikkinen E, Heikkinen R-L, Ruoppila I, (eds). Functional capacity and health of elderly people – the Evergreen Project. *Scand J Soc Med Suppl*, 53, 1–18.

HEINONEN, O. P. 1966. Autoklinikka. *Duodecim,* 82,1161-1164 (in Finnish).

HEISTARO, S. (ed.) 2008. Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008. http://urn.fi/URN:NBN:fi-fe201204193320

HEISTARO, S. & KOSKINEN, S. 2008. Resurvey of persons who participated in the Mini-Finland Survey. In: HEISTARO, S. (ed.) Methodology report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 150. http://urn.fi/URN:NBN:fi-fe201204193320

HELAKORPI, S., UUTELA, A., PRÄTTÄLÄ, R. & PUSKA, P. 2000. Suomalaisten aikuisväestön terveyskäyttäytyminen ja terveys, kevät 2000. Kansanterveyslaitoksen julkaisuja KTL/B8. Helsinki (in Finnish).

HELIÖVAARA, M. 2008. Spirometry. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 66-68. http://urn.fi/URN:NBN:fi-fe201204193320

HELLDÁN, A., RAULIO, S., KOSOLA, M., TAPANAINEN, H., OVASKAINEN, M.-L.-. & VIRTANEN, S. 2013. Finravinto 2012 -tutkimus – The National FINDIET 2012 Survey. Raportti 16/2013. Helsinki.

HOPPENFELD, S. 1976. Physical examination of the spine and extremities. Pearson Education. New York.

HU G, TUOMILEHTO J, SILVENTOINEN K, BARENGO NC, PELTONEN M & JOUSILAHTI P 2005. The effects of physical activity and body mass index on cardiovascular, cancer and all-cause mortality among 47 212 middle-aged Finnish men and women. *Int J Obes,* 29**,** 894-902.

HUSU, P., PARONEN, O., SUNI, J. & VASANKARI, T. 2011. Suomalaisten fyysinen aktiivisuus ja kunto 2010. Terveyttä edistävän liikunnan nykytila ja muutokset. Opetus- ja kulttuuriministeriön julkaisuja 2011:15.

HÄNNINEN, T., PULLIAINEN, V., SALO, J., HOKKANEN, L., ERKINJUNTTI, T., KOIVISTO, K., VIRAMO, P., SOININEN, H. & SUOMEN MUISTITUTKIMUSYKSIKÖIDEN ASIANTUNTIJARYHMÄ 1999. Kognitiiviset testit muistihäiriöiden ja alkavan dementian varhaisdiagnostiikassa: CERAD-tehtäväsarja. *Suom Lääkäril, 15, 1967–75* (in Finnish).

HÄRKÄNEN, T., KARVANEN, J., TOLONEN, H., LEHTONEN, R., DJERF, K., JUNTUNEN, T. & KOSKINEN, S. 2016. Systematic handling of missing data in complex study designs - experiences from the Health 2000 and 2011 Surveys. *J Appl Stat* http://dx.doi.org/10.1080/02664763.2016.1144725.

JACOBI, F., WITTCHEN, H. U., HOLTING, C., HOFLER, M., PFISTER, H., MULLER, N. & LIEB, R. 2004. Prevalence, co-morbidity and correlates of mental disorders in the general population: results from the German Health Interview and Examination Survey (GHS). *Psychol Med,* 34**,** 597-611.

JAFFRIN, M. Y. 2009. Body composition determination by bioimpedance: an update. *Curr Opin Clin Nutr Metab Care,* 12**,** 482-6.

JOUKAMAA, M., LEHTINEN, V., KARLSSON, H. & ROUHE, E. 1994. SCL-25 and recognition of mental disorders reported by primary health care physicians. *Acta Psychiatr Scand,* 89**,** 329-3.

JYLHÄ, M. & ARO, S. 1989. Social ties and survival among the elderly in Tampere, Finland. *Int J Epidemiol,* 18, 158-164.

KAARTINEN, N. E., TAPANAINEN, H., VALSTA, L. M., SIMILÄ, M. E., REINIVUO, H., KORHONEN, T., HARALD, K., J.G., E., PELTONEN, M. & MÄNNISTÖ, S. 2011. Relative validity of a FFQ in measuring carbohydrate fractions, dietary glycaemic index and load: exploring the effects of subject characteristics. *Br J Nutr,* 107, 1367-75.

KALIMO, R., HAKANEN, J. & TOPPINEN-TANNER, S. 2006. Maslachin yleinen työuupumuksen arviointimenetelmä MBI-GS. [The Finnish version of Maslach's Burnout Inventory - General Survey]. Työterveyslaitos. Helsinki.

KALLELA, M., WESSMAN, M. & FÄRKKILÄ, M. 2001. Validation of a migraine-specific questionnaire for use in family studies. *Eur J Neurol,* 8**,** 61-6.

KARASEK, R., BRISSON, C., KAWAKAMI, N., HOUTMAN, I., BONGERS, P. & AMICK, B. 1998. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *J Occup Health Psychol,* 3, 322-55.

KATZ, S., DOWNS, T. D., CASH, H. R. & GROTZ, R. C. 1970. Progress in development of the index of ADL. *Gerontologist,* 10, 20-30.

KATZ, S., FORD, A. B., MOSKOWITZ, R. W., JACKSON, B. A. & JAFFE, M. W. 1963. Studies of illness in the aged. The index of ADL: a standardized measure of biologocal and psychosocial function. *JAMA,* 185, 914-19.

KESSLER, R. C., ANGERMEYER, M., ANTHONY, J. C., R, D. E. G., DEMYTTENAERE, K., GASQUET, I., G, D. E. G., GLUZMAN, S., GUREJE, O., HARO, J. M., KAWAKAMI, N., KARAM, A., LEVINSON, D., MEDINA MORA, M. E., OAKLEY BROWNE, M. A., POSADA-VILLA, J., STEIN, D. J., ADLEY TSANG, C. H., AGUILAR-GAXIOLA, S., ALONSO, J., LEE, S., HEERINGA, S., PENNELL, B. E., BERGLUND, P., GRUBER, M. J., PETUKHOVA, M., CHATTERJI, S. & USTUN, T. B. 2007. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry,* 6**,** 168-76.

KESSLER, R. C. & USTUN, T. B. 2004. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res,* 13**,** 93-121.

KODAMA, S., SAITO, K. & TANAKA, S. 2009. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women – a meta-analysis. *JAMA*, 301, 2024-25.

KOLU, P., VASANKARI, T. & LUOTO, R. 2014. Liikkumattomuus ja terveydenhuollon kustannukset. *Lääkärilehti,* 12**,** 885-889.

KOSKINEN, S., LUNDQVIST, A. & RISTILUOMA, N. (eds.) 2012. Terveys, toimintakyky ja hyvinvointi Suomessa 2011. Raportti 68/2012. Tampere. http://urn.fi/URN:ISBN:978-952-245-769-1

KOSKINEN, S., SAINIO, P., SALMINEN, I. & REUNANEN, A. 2008. Home health examination. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 126-31. http://urn.fi/URN:NBN:fi-fe201204193320.

KÄYPAHOITOSUOSITUS. 2015. Alkoholiongelmaisen hoito. http://www.kaypahoito.fi/web/kh/suositukset/suositus?id=hoi50028

LAIHO, J. 2004. Katoanalyysi. In: LAIHO, J. & NIEMINEN, T. (eds.) Terveys 2000 -tutkimus. Aikuisväestön haastatteluaineiston tilastollinen laatu. Otanta-asetelma, tiedonkeruu, vastauskato ja estimointi- ja analyysiasetelma. Tilastokeskus, tutkimuksia 239. Helsinki. 28-47 (in Finnish).

LAIHO, J., DJERF, K. & LEHTONEN, R. 2008. Sampling design. In: HEISTARO, S. (ed.) Methodology report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 13. http://urn.fi/URN:NBN:fi-fe201204193320.

LAWTON, M. P. & BRODY, E. M. 1969. Assessment of older people: self-maintaining and instrumental activities ofdaily living. *Gerontologist,* 9, 179-86.

LEE, J. 1981. Covariance adjustment of rates based on the multiple logistic regression model. *J Chronic Dis,* 34**,** 415−26.

LEHTO, A.-M. 1991. Työelämän laatu ja tasa-arvo tuotannossa. [The quality of working life and equality in production.] Statistics Finland. Helsinki.

LEHTONEN, R. & PAHKINEN, E. 2004. Practical methods for design and analysis of complex surveys. Revised 2nd Edition. John Wiley & Sons. Chichester.

LEONARDI, M., CHATTERJI, S., KOSKINEN, S., AYUSO-MATEOS, J. L., HARO, J. M., FRISONI, G., FRATTURA, L., MARTINUZZI, A., TOBIASZ-ADAMCZYK, B., GMUREK, M., SERRAANO, R. & FINOCCHIARO, C. 2014. Determinants of Health and Disability in Ageing Population: The COURAGE in Europe Project (Collaborative Research on Ageing in Europe). *Clin Psychol Psychother,* 21, 193–198.

LUDBROOK, J. 2002. Statistical techniques for comparing measurers and methods of measurement: a critical review. *Clin Exp Pharmacol Physiol,* 29, 527-36.

LUMLEY, T. 2004. Analysis of complex survey samples. *Journal of Statistical Software,* 9**,** 1–19.

LYNN, P. 2009. Methodology of Longitudinal Surveys*.* Wiley. Chippenham.

LÄHTEENMÄKI, S., AALTO-SETÄLÄ, T., SUOKAS, J. T., SAARNI, S. E., PERÄLÄ, J., SAARNI, S. I., ARO, H., LÖNNQVIST, J. & SUVISAARI, J. M. 2009. Validation of the Finnish version of the SCOFF questionnaire among young adults aged 20 to 35 years. *BMC Psychiatry,* 9**,** 5.

MASLACH, C., JACKSON, S.E., LEITER, M.P. 1996. The Maslach Burnout Inventory (3rd ed.). Consulting Psychologists Press. Palo Alto, CA.

MATTILA, A. K., SALMINEN, J. K., NUMMI, T. & JOUKAMAA, M. 2006. Age is strongly associated with alexithymia in the general population. *J Psychosom Res,* 61**,** 629-35.

MCWHINNIE, J. R. 1981. Disability assessment in population surveys: results of the OP.E.C.D. Common Development Effort. *Rev Epidemiol Santé Publique,* 29, 413-19.

MOLENBERGHS, G. & KENWARD, M. G. 2007. Missing Data in Clinical Studies. Wiley. Cornwall.

MORGAN, J. F., REID, F. & LACEY, J. H. 1999. The SCOFF questionnaire: assessment of a new screening tool for eating disorders. *BMJ,* 319**,** 1467-8.

MORRIS, J. C., HEYMAN, A., MOHS, R. C., HUGHES, J. P., VAN BELLE, G., FILLENBAUM, G., MELLITS, E. D. & CLARK, C. 1989. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's Disease. *Neurology, 39, 1159–65*.

MÄNNISTÖ, S., VIRTANEN, M., MIKKONEN, T. & PIETINEN, P. 1996. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. *J Clin Epidemiol,* 49**,** 401-409.

MÄNTY, M., HEINONEN, A., LEINONEN, R., TÖRMÄKANGAS, T., SAKARI-RANTALA, R., HIRVENSALO, M., VON BONSDORFF, M. B. & RANTANEN, T. 2007. Construct and predictive validity of a self-reported measure of preclinical mobility limitation. *Arch Phys Med Rehabil,* 88**,** 1108-13.

NELSON, T. L., PALMER, R. F. & PEDERSEN, N. L. 2004. The metabolic syndrome mediates the relationship between cynical hostility and cardiovascular disease. *Exp Aging Res,* 30, 163-77.

NIEMINEN, T. 2015. Healthier together? Social capital, health behavior and health. Academic dissertation, University of Helsinki, Faculty of Medicine.

NIEMINEN, T., MARTELIN, T., KOSKINEN, S., SIMPURA, J., ALANEN, E., HÄRKÄNEN, T. & AROMAA, A. 2008. Measurement and socio-demographic variation of social capital in a large population-based survey. *Soc Indic Res,* 85, 405–23.

OJA, P. & TUXWORTH, W. (eds.) 1995. Eurofit for adults. Assessment of health-related fitness. Finland. Council of Europe Publishing.

PAALANEN, L., MÄNNISTÖ, S., VIRTANEN, M. J., KNEKT, P., RÄSÄNEN, L., MONTONEN, J. & PIETINEN, P. 2006. Validity of a food frequency questionnaire varied by age and body mass index. *J Clin Epidemiol,* 59**,** 994-1001.

PIETINEN, P., HARTMAN, A. M., HAAPA, E., RÄSÄNEN, L., HAAPAKOSKI, J., PALMGREN, J., ALBANES, J., ALBANES, D., VIRTAMO, J. & HUTTUNEN, J. K. 1988. Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. *Am J Epidemiol,* 128**,** 655-66.

PIIRAINEN, H., ELO, A. L., HIRVONEN, M., KAUPPINEN, K., KETOLA, R., LAITINEN, H., LINDSTRÖM, K., REIJULA, K., RIALA, R., VILUKSELA, M. & VIRTANEN, S. 2000. Työ ja terveys -haastattelututkimus v. 2000. Taulukkoraportti. Työterveyslaitos, Helsinki (in Finnish).

PIRKOLA, S., LÖNNQVIST, J. & MENTAL HEALTH TEAM 2008. Mental health interview. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 111-14. http://urn.fi/URN:NBN:fi-fe201204193320.

POWER, M. 2003. Development of a common instrument for quality of life. In: NOSIKOV, A. & GUDEX, C. EUROHIS: Developing common instruments for health surveys. *Biomedical and Health Research*, 57, 145-163.

PULKKI, L., KIVIMÄKI, M., ELOVAINIO, M., J., V. & KELTIKANGAS-JARVINEN, L. 2003. Contribution of socioeconomic status to the association between hostility and cardiovascular risk behaviors: a prospective cohort study. *Am J Epidemiol,* 158, 736-42.

PULLIAINEN, V., HOKKANEN, L., SALO, J. & HÄNNINEN, T. 1999. CERAD, kognitiivinen tehtäväsarja, käsikirja. Suomen Alzheimer –tutkimusseura. Kuopio (in Finnish).

PUROLA, T., KALIMO, E., SIEVERS, K. & NYMAN, K. 1967. Sairastavuus ja lääkintäpalvelusten käyttö Suomessa ennen sairausvakuutusta. Kansaneläkelaitoksen julkaisuja A:1. Helsinki.

QUINTAS, R., RAGGI, A., BUCCIARELLI, P., FRANCO, M. G., ANDREOTTI, A., CABALLERO, F. F., OLAYA, B., CHATTERJI, S., GALAS, A., MERILÄINEN-PORRAS, S., FRISONI, G. B., RUSSO, E., MINICUCI, N., POWER, M. & LEONARDI, M. 2014. The COURAGE Built Environment Outdoor Checklist: An Objective Built Environment Instrument to Investigate the Impact of the Environment on Health and Disability. *Clin Psychol Psychother,* 21, 204–14.

R CORE TEAM 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna.

RABIN, R. & DE CHARRO, F. 2001. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med,* 33, 337-43.

RAGGI, A., QUINTAS, R., BUCCIARELLI, P., FRANCO, M. G., ANDREOTTI, A., MIRET, M., ZAWISZA, K., OYALA, B., CHATTERJI, S., SAINIO, P., FRISONI, G. B., MARTINUZZI, A., MINICUCI, N., POWER, M. & LEONARDI, M. 2014. Validation of the COURAGE Built Environment Self-Reported Questionnaire. *Clin Psychol Psychother,* 21, 215–26.

REINIVUO, H., HIRVONEN, T., OVASKAINEN, M. L., KORHONEN, T. & VALSTA, L. M. 2010. Dietary survey methodology of FINDIET 2007 with a risk assessment perspective. *Public Health Nutr,* 13**,** 915-9.

RESEARCH TRIANGLE INSTITUTE 2008. SUDAAN Language Manual, Release 10.0. Research Triangle Park NC. Research Triangle Institute.

REUNANEN, A. 2008. Blood pressure and heart rate. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 60-61. http://urn.fi/URN:NBN:fi-fe201204193320.

REUNANEN, A. & HELIÖVAARA, M. 2008. Symptoms interview. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 56-59. http://urn.fi/URN:NBN:fi-fe201204193320.

RINTAMÄKI, R., GRIMALDI, S., ENGLUND, A., HAUKKA, J., PARTONEN, T., REUNANEN, A., AROMAA, A. & LÖNNQVIST, J. 2008. Seasonal changes in mood and behavior are linked to metabolic syndrome. *PLoS One,* 3**,** e1482.

ROBINS, J. M., ROTNITZKY, A. & ZHAO, L. P. 1994. Estimation of regression coefficients when some regressors are not always observed. *J Am Stat Assoc, 89, 846–866*.

ROSE, G. A. & BLACKBURN, H. 1968. Cardiovascular survey methods. World Health Organization. Geneva.

ROSE, G. A., BLACKBURN, H., GILLUM, R. F. & PRINEAS, R. J. 1982. Cardiovascular survey methods. Second edition. World Health Organization. Geneva.

ROSENTHAL, N. E., BRADT, G. H. & WEHR, T. A. 1984. Seasonal Pattern Assessment Questionnaire. National Institute of Mental Health. Bethesda (MD).

ROSOW, I. & BRESLAU, N. 1966. A Guttman health scale for the aged. *J Gerontol,* 21, 556–9.

RUBIN, D. B. 1987. Multiple imputation for nonresponse in surveys. Wiley. New York.

RUDANKO, S. L. & KOSKINEN, S. 2008. Study of visual acuity. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 93-94. http://urn.fi/URN:NBN:fi-fe201204193320.

SAINIO, P. & MALMBERG, J. 2008. Tests of physical capacity administered to subjects aged 55 or over. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 99. http://urn.fi/URN:NBN:fi-fe201204193320.

SALTIN, B. & GRIMBY, G. 1968. Physiological Analysis of Middle-Aged and Old Former Athletes. *Circulation,* XXXVIII**,** 1104-1115.

SAMMONG PRESTON ROLYAN 2003. Jamar Adjustable Hand Dynamometer Owener's Manual. Sammong Preston Rolyan. Grand Rapids, MI.

SARASON, I. G., LEVINE, H. M., BASHAM, R. B. & SARASON, B. R. 1983. Assessing social support: The social support questionnaire. *J Pers Soc Psychol,* 44, 127–139.

SAS INSTITUTE INC 2010. SAS/STAT® 9.22 User's Guide. SAS Institute Inc. Cary, NC.

SCHAFER, J. L. 1999. Multiple imputation: a primer. *Stat Methods Med Res*, 8**,** 3-15

SCHWARZ, G. 1978. Estimating the dimension of a model. *The Annals of Statistics,* 6**,** 461-464.

SIEVERS, K., MELKAS, T. & HELIÖVAARA, M. 1985. Mini-Suomi -terveystutkimuksen toteutus. Osa 3. Tuki- ja liikuntaelinten sairauksien tutkimusmenetelmät. Kansaneläkelaitoksen julkaisuja ML:50. Helsinki ja Turku (in Finnish).

SIMONSICK, E. M., NEWMAN, A. B., NEVITT, M. C., KRITCHEVSKY, S. B., FERRUCCI, L., GURALNIK, J. M. & HARRIS, T. 2001. Measuring higher level physical function in well-functioning older adults: expanding familiar approaches in the Health ABC study. *J Gerontol A Biol Sci Med Sci,* 56**,** M644-9.

SINTONEN, H. 1994. The 15D Measure of Health Related Quality of Life: Reliability, Validity and Sensitivity of its Health State Descriptive System. National Centre for Health Program Evaluation, Working Paper 41, Melbourne.

SINTONEN, H. 1995. The 15D Measure of Health Related Quality of Life. II Feasibility, Reliability and Validity of its Valuation System. National Centre for Health Program Evaluation, Working Paper 42, Melbourne.

SLADE, G. D. 1997. Derivation and validation of a short-form oral health impact profile. *Community Dent Oral Epidemiol*, 25, 284-90.

SLADE, G. D. & SPENCER, A. J. 1994. Development and evaluation of the Oral Health Impact Profile. *Community Dental Health,* 11, 3-11.

SOLOVIEVA, S., PEHKONEN, I., KAUSTO, J., MIRANDA, H., SHIRI, R., KAUPPINEN, T., HELIÖVAARA, M., BURDORF, A., HUSGAFVEL-PURSIAINEN, K. & VIIKARI-JUNTURA, E. 2012. Development and validation of a job exposure matrix for physical risk factors in low back pain. *PLoS One,* 7, e48680.

SOSIAALI- JA TERVEYSMINISTERIÖ. 2015. Istu vähemmän – Voi paremmin! Kansalliset suositukset istumisen vähentämiseen. http://urn.fi/URN:ISBN:978-952-00-3726-0

STATACORP 2009. Stata: Release 11. Statistical Software. College Station, TX: StataCorp LP.

STENHOLM, S., HÄRKÄNEN, T., SAINIO, P., HELIÖVAARA, M. & KOSKINEN, S. 2012. Long-term changes in handgrip strength in men and women–accounting the effect of right censoring due to death. *J Gerontol A Biol Sci Med Sci,* 67**,** 1068-74.

SUNI, J. H., OJA, P. & LAUKKANEN, R. T. 1996. Health-related fitness test battery for adults: aspects of reliability. *Arch Phys Med Rehabil,* 77**,** 399–405.

SUOMINEN-TAIPALE, A. L., NORDBLAD, A., VEHKALAHTI, M. & AROMAA, A. (eds.) 2008. Oral Health in the Finnish adult population. Health 2000 survey. Publications of the National Public Health Institute B 25/2008. http://urn.fi/URN:ISBN:978-951-740-852-3.

SUUTAMA, T., SULKAVA, R. & KOSKINEN, S. 2008. Tests of congnitive functioning. In: HEISTARO, S. (ed.) Methodology Report. Health 2000 Survey. Publications of the National Public Health Institute B26/2008, 95. http://urn.fi/URN:NBN:fi-fe201204193320.

THOMAS, B. J., CORNISH, B. H., PATTEMORE, M. J., JACOBS, M. & WARD, L. C. 2003. A comparison of the whole-body and segmental methodologies of bioimpedance analysis. *Acta Diabetol,* 40**,** 236-7.

TIIKKAINEN, P. & HEIKKINEN, R. L. 2011. Sosiaalisen toimintakyvyn arviointi ja mittaaminen väestötutkimuksissa (Assessment of social functioning in population surveys). TOIMIA-suositus. Terveyden ja hyvinvoinnin laitos. http://www.thl.fi/toimia/tietokanta/suositus/18/ (in Finnish).

TINDLE, H. A., CHANG, Y. F., KULLER, L. H., MANSON, J. E., ROBINSON, J. G., ROSAL, M. C., SIEGLE, G. J. & MATTHEWS, K. A. 2009. Optimism, cynical hostility, and incident coronary heart disease and mortality in the Women's Health Initiative. *Circulation,* 120, 656-62.

TOLONEN, H. 2013. EHES Manual - Part A - Planning and Preparation of the Survey. THL Directions 2013_001. http://urn.fi/URN:ISBN:978-952-245-842-1.

TUOMI, K., ILMARINEN, J., JAHKOLA, M., KATAJARINNE, L. & TULKKI, A. 2006. Work Ability Index. First reprint of the 2nd revised edition. Finnish Institute of Occupational Health. Helsinki.

UKK INSTITUTE. 2009. Physical activity pie. Finnish guidelines for health-enhancing physical activity for adults aged 18 to 64 years (English version). http://www.ukkinstituutti.fi/en/products/physical_activity_pie

VAN BUUREN, S., BOSHUIZEN, H. C. & KNOOK, D. L. 1999. Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in medicine,* 18, 681-94.

WARD, L. C. 2012. Segmental bioelectrical impedance analysis: an update. *Curr Opin Clin Nutr Metab Care,* 15**,** 424-9.

VEHKALAHTI, M., PAUNIO, I., NYYSSÖNEN, V. & AROMAA, A. (eds.) 1991. Suomalaisten aikuisten suunterveys ja siihen vaikuttavat tekijät (Oral health and factors related to it in the Finnish adult population). Kansaneläkelaitoksen julkaisuja. AL: 34. Helsinki ja Turku (in Finnish; Abstract in English).

VEIJOLA, J., JOKELAINEN, J., LAKSY, K., KANTOJARVI, L., KOKKONEN, P., JARVELIN, M. R. & JOUKAMAA, M. 2003. The Hopkins Symptom Checklist-25 in screening DSM-III-R axis-I disorders. *Nord J Psychiatry,* 57**,** 119-23.

WEISS, R. 1973. Loneliness. The experience of emotional and social isolation. MIT Press. Cambridge.

WELIN, L., BRESATER, L. E., ERIKSSON, H., HANSSON, P. O., WELIN, C. & ROSENGREN, A. 2003. Insulin resistance and other risk factors for coronary heart disease in elderly men. The Study of Men Born in 1913 and 1923. *Eur J Cardiovasc Prev Rehabil,* 10**,** 283-8.

WHO, 2002. Towards a Common Language for Functioning, Disability and Health: ICF. Geneva.

WHO, 2010. Global recommendations on physical activity for health.

VIIKARI-JUNTURA, E., RAUAS, S., MARTIKAINEN, R., KUOSMA, E., RIIHIMÄKI, H., TAKALA, E. P. & SAARENMAA, K. 1996. Validity of self-reported physical work load in epidemiologic studies on musculoskeletal disorders. *Scand J Work Environ Health,* 22, 251-9.

WILHELMSEN, L., WEDEL, H. & TIBBLIN, G. 1973. Multivariate analysis of risk factors for coronary heart disease. *Circulation,* 48**,** 950-8.

WILLETT, W. 2013. Nutritional Epidemiology (third edition). Oxford University Press. New York.

WINER, B. J. 1971. Statistical principles in experimental design. 2nd ed. McGraw-Hill and Kogakusha. Tokyo.

WIRTH, K. E., TCHETGEN, E. J. & MURRAY, M. 2010. Adjustment for Missing Data in Complex Surveys Using Doubly Robust Estimation: Application to Commercial Sexual Contact Among Indian Men. *Epidemiology,* 21, 863–71.

WITTCHEN, H. U., LACHNER, G., WUNDERLICH, U. & PFISTER, H. 1998. Test-retest reliability of the computerized DSM-IV version of the Munich-Composite International Diagnostic Interview (M-CIDI). *Soc Psychiatry Psychiatr Epidemiol,* 33**,** 568-78.

ZAWISZA, K., GALAS, A., TOBIASZ-ADAMCZYK, B., CHATTERJI, S., HARO, J. M., MIRET, M., KOSKINEN, S., POWER, M. & LEONARDI, M. 2014. The Validity of the Instrument to Evaluate Social Network in the Ageing Population: The Collaborative Research on Ageing in Europe Social Network Index. *Clin Psychol Psychother,* 21, 227–41.

APPENDICES

APPENDIX 1. Health 2011 organisation during the planning and data collection phases

Advisory board

Pekka Puska (chair), Marja Vaarama (vice chair), Tomi Mäki-Opas (secretary), Olli Kangas, Mikko Kautto, Arto Koho, Seppo Koskinen, Riitta-Liisa Lappetelainen, Vesa Pohjola, Timo Silvola, Tommi Vasankari and Eira Viikari-Juntura

Steering committee

Marja Vaarama (chair), Seppo Koskinen (vice chair), Tomi Mäki-Opas (secretary), Marina Erhola, Juhani Eskola, Elina Kestilä-Kekkonen/Matti Sarjakoski, Marja Lampola, Annamari Lundqvist, Mikko Nissinen, Ritva Prättälä /Eila Linnanmäki and Erkki Vartiainen

Management group (former project group)

Seppo Koskinen (chair), Pirkko Alha / Ulla Laitinen (secretary), the chair and vice-chair of the expert groups (see below)

Core management team

Seppo Koskinen (chair), Ulla Laitinen (secretary), Pirkko Alha, Paul Knekt, Päivikki Koponen, Annamari Lundqvist, Tomi Mäki-Opas and Noora Ristiluoma

Scientific advisory board (TAR)

Seppo Koskinen (chair), Pirkko Alha / Noora Ristiluoma (secretary), Markku Heliövaara, Tommi Härkänen, Paul Knekt, Päivikki Koponen, Maarit Laaksonen, Annamari Lundqvist and Tomi Mäki-Opas

Expert groups

Living conditions and health. Sakari Karvonen (chair), Katri Hannikainen-Ingman (secretary), Jenni Blomgren, Timo Kauppinen, Laura Kestilä, Eero Lahelma, Tuija Martelin, Pasi Moisio and Tiina Pensola

Health behaviours (subgroups: F=Food, P=Physical activity, A=Alcohol & Smoking, S=Sleep). Ritva Prättälä (chair), Tomi Mäki-Opas (secretary), Hannu Alho, Katja Borodulin, Iris Erlund, Ari Haukkala, Antero Heloma, Pauliina Husu, Mikko Harma, Tuija Jääskeläinen (F secretary), Niina Kaartinen, Noora Kanerva, Erkki Kronholm (S chair), Marjaana Lahti-Koski, Annamari Lundqvist, Pia Mäkelä (A chair), Satu Männistö (F chair), Laura Paalanen, Susanna Raulio, Eva Roos, Päivi Sainio, Laura Sares-Jäske, Harri Sievänen, Sari Stenholm, Jaana Suni, Sakari Suominen, Katri Sääksjärvi, Heli Valkeinen, Tommi Vasankari (P chair) and Miira Vehkalahti

Obesity and body composition. Annamari Lundqvist (chair), Satu Männistö (vice chair), Laura Sares-Jäske (secretary), Markku Heliövaara, Antti Jula, Paul Knekt, Seppo Koskinen, Marjaana Lahti-Koski, Susanna Raulio, Aila Rissanen, Suoma Saarni, Veikko Salomaa and Sari Stenholm

Cardiovascular diseases and diabetes. Antti Jula (chair), Veikko Salomaa (vice chair), Eija Viholainen (secretary), Matti Jauhiainen, Risto Kaaja, Antero Kesäniemi, Johanna Kuusisto, Mika Kähönen, Markku Laakso, Jaana Leiviskä, Markku Nieminen, Teemu Niiranen, Lasse Oikarinen, Janne Rapola, Jaakko Tuomilehto and Olavi Ukkola

Respiratory diseases and allergies. Tuula Vasankari (chair), Merja Kanervisto (secretary), Markku Heliövaara, Pekka Jousilahti, Tarja Laitinen, Tiina Mattila and Seppo Saarelainen

Musculoskeletal diseases. Eira Viikari-Juntura (chair), Markku Heliövaara (vice chair), Leena Kaila-Kangas (secretary), Jari Arokoski, Kirsti Husgafvel-Pursiainen, Jaro Karppinen, Heikki Kröger, Päivi Leino-Arjas, Rahman Shiri, Svetlana Solovieva and Esa-Pekka Takala

Mental health. Jaana Suvisaari (chair), Satu Viertiö (secretary), Mauri Aalto, Kirsi Ahola, Olli Kiviruusu, Olavi Lindfors, Aino Mattila, Niina Markkula, Timo Partonen, Sebastián Pena Fajuri, Jonna Perälä, Sami Pirkola, Samuli Saarni, Suoma Saarni and Annamari Tuulio-Henriksson

Oral health. Liisa Suominen (chair), Sinikka Varsio (vice chair), Sari Helminen, Matti Knuuttila, Satu Lahti, Anne Nordblad, Ulla Tyyni and Miira Vehkalahti

Reproductive health. Riitta Luoto (chair), Päivikki Koponen (vice chair), Pirkko Alha (secretary), Elina Hemminki, Risto Kaaja, Reija Klemetti and Heljä-Maria Surcel

Cancer. Eero Pukkala (chair), Paul Knekt (vice chair), Katri Sääksjärvi (secretary), Lauri Aaltonen, Mika Gissler, Markku Heliövaara, Tommi Härkänen, Jaakko Kaprio, Timo Kauppinen, Ilmo Keskimäki, Matti Lehtinen, Esa Läärä, Nea Malila, Jukka Marniemi, Pekka Martikainen, Satu Männistö, Suoma Saarni, Risto Sankila, Harri Vainio and Jarmo Virtamo

Infectious diseases. Petri Ruutu (chair), Henrikki Brummer-Korvenkontio, Irja Davidkin, Eija Hiltunen-Back, Pia Kivelä, Markku Kuusi, Kirsi Liitsola, Mulki Mölsä, Hanna Nohynek, Matti Ristola, Mika Salminen, Anja Siitonen, Heljä-Marja Surcel and Paula Tiittala

Functioning, working capacity, need for help and rehabilitation. Päivi Sainio (chair), Shadia Rask (secretary), Jenni Blomgren, Pertti Era, Raija Gould, Tuula Hurnasti, Pauliina Husu, Seppo Koskinen, Erkki Kronholm, Arja Laitinen, Minna-Liisa Luoma, Matti Mäkelä, Tiina Pensola, Taina Rantanen, Noora Ristiluoma, Sari Stenholm, Raimo Sulkava, Jaana Suni, Pirjo Tiikkainen, Annamari Tuulio-Henriksson, Mariitta Vaara and Heli Valkeinen

Use and need for health services. Ilmo Keskimäki (chair), Kristiina Manderbacka (vice chair), Pirkko Alha (secretary), Unto Häkkinen, Sari Kehusmaa, Päivikki Koponen, Olavi Lindfors, Jaana Martikainen, Lien Nguyen, Monica Röberg and Riitta Sauni

Genetics. Veikko Salomaa (chair), Markku Heliövaara, Kirsti Husgafvel-Pursiainen, Antti Jula, Seppo Koskinen, Markus Perola, Samuli Ripatti and Jaana Suvisaari

Laboratory. Jouko Sundvall (chair), Jaana Leiviskä (secretary), Pirkko Alha, Päivi Laiho, Jukka Lauronen, Britt-Marie Loo and Laura Lund

Statistics. Tommi Härkänen (chair), Kari Djerf, Paul Knekt, Risto Lehtonen and Esa Virtala

Quality of data. Esa Virtala (chair), Pirkko Alha (vice chair), Harri Rissanen (vice chair), Tuija Jääskeläinen (secretary), Paul Knekt, Sirkka Rinne, Noora Ristiluoma and Laura Sares-Jäske

Co-ordinating group of the fieldwork. Tomi Mäki-Opas (chair), Ulla Laitinen (secretary), Pirkko Alha, Minna Kuurne-Koivisto, Hannele Lehtonen, Laura Lund, Annamari Lundqvist, Satu Merilainen-Porras, Sirkka Rinne, Noora Ristiluoma, Tarja Räty and Ulla Tyyni

Persons in charge of the different study phases

Seppo Koskinen, Arpo Aromaa, Sebastián Peña, Päivikki Koponen, Annamari Lundqvist and Pirkko Alha were in charge of the planning and execution of Health 2011 Survey together with several expert groups

Seppo Koskinen (chair), Tomi Mäki-Opas (vice chair), Ulla Laitinen (secretary), Arpo Aromaa, Pirkko Alha, Markku Heliövaara, Paul Knekt, Päivikki Koponen, Sebastián Peña, Annamari Lundqvist and Tarja Räty in charge of the planning and execution of Health 2011 Survey together with several expert groups

Tommi Härkänen and Sirkka Rinne were in charge of the sample design and Tommi Härkänen calculated the sample weights.

Pirkko Alha, Tarja Räty and Ulla Tyyni were in charge of recruiting the fieldwork personnel. Pirkko Alha and Tarja Räty interviewed the members of the Helsinki fieldwork team. Ulla Tyyni and Helena Holopainen interviewed the members of Kuopio and Oulu fieldwork teams. Pirkko Alha, Päivikki Koponen and Satu Meriläinen-Porras interviewed the members of Turku fieldwork team. Courage-nurses were recruited by using similar protocol and Satu Meriläinen-Porras participated in the interviews. Liisa Suominen and Anne Nordblad recruited the dentists.

Tomi Mäki-Opas, Pirkko Alha and Päivikki Koponen were in charge of planning the training of the fieldwork personnel. Moreover, experts of the measurements participated in the planning and training. The training of the fieldwork was coordinated by Helmi Koskinen and Ulla Laitinen.

Paul Knekt, Arpo Aromaa, Päivikki Koponen, Markku Heliövaara and Sirkka Rinne were in charge of the planning and execution of the quality assurance during the fieldwork.

Tomi Mäki-Opas, Eeva-Liisa Kallonen, Tarja Räty, Mikko Pekkarinen, Lea Kurki, Elisa Kostiainen, Satu Meriläinen-Porras and Eeva Parviainen were in charge of the communications before and during the fieldwork.

Research facilities were acquired by Tomi Mäki-Opas, Ulla Tyyni and Seppo Koskinen. Moreover, Ulla Tyyni and Tomi Mäki-Opas planned the time schedules for the fieldwork teams.

Health examinations were conducted by five fieldwork teams, which included 87 fieldwork nurses in total. Fieldwork teams were led by five head nurses: Kimmo Porras (Helsinki), Elina

Poutanen (Tampere), Marjatta Härkäs (Turku), Ulla Tyyni (Kuopio) and Liisa Moilanen (Oulu). Tomi Mäki-Opas operated as the fieldwork manager.

The COURAGE sub-study was planned, organized and coordinated by Satu Meriläinen-Porras, Minna Kuurne-Koivisto, Ulla Laitinen and Seppo Koskinen and executed together with Health 2011 Survey. The COURAGE fieldwork was conducted by 14 nurses.

Persons in charge of the health examination were Noora Ristiluoma (appointments), Sirkka Rinne (registration), Ulla Laitinen (measurements), Liisa Suominen (oral health), Laura Lund (laboratory), Päivi Sainio (functioning), Jaana Suvisaari (mental health), Päivikki Koponen and Seppo Koskinen (interviews), Pirkko Alha (final check-up), Heli Valkeinen and Jaana Suni (physical activity measurements), Ulla Laitinen, Päivi Sainio (COURAGE) and Annamari Lundqvist (complementary data collection).

Appointments were handled at THL by Veera Jussila, Varpu Mäkelä, Satu Suihko and Sari Ullgren-Lajunen. Ritva Topp and Juho Sipilä also assisted with the study.

Laboratory analyses were conducted at THL, TLAB, under the leadership of Jouko Sundvall. Laura Lund was in charge of training the fieldwork personnel for sampling of the blood and quality control during the fieldwork.

The complementary data collection was planned, organized and lead by Annamari Lundqvist. IT and logistics design was planned and conducted by Teijo Kalliomäki, Tarja Räty, Ulla Tyyni, Mikko Pekkarinen, Marko Grönholm, Jaason Haapakoski, Talvikki Leinonen, Tuomas Karhu and Timo Lamminjoki. Computer aided interview forms were conduted by Vesa Tanskanen and Juhani Mäki. The webpages for the Health 2011 were designed by Mikko Pekkarinen.

Employment issues were handled by Virpi Killström, billing by Terttu Malkamäki and travel issues by Jaana Lindh and Merja Martenson.

Several people at THL's administration, in the other organisations outside THL as well as administrations of various study-sites contributed significantly to the planning stage and carrying out the Health 2011 Survey.

APPENDIX 2. Fieldwork personnel in the Health 2011 Survey

Helsinki

Field nurses: *Kimmo Porras (head nurse), Helena Andersson, Minna-Liisa Collins-Cona, Valentina Daderchenkova, Emma Dufva, Susanna Helenius, Kati Immonen, Sonja Jatuli, Merja Joutsenlahti, Eini Oinonen, Pirjo Rahmonen, Suvi-Päivikki Salo, Anna Sarpio, Tiina Saukkonen, Sirpa Sillanpää-Savolainen, Heli Sipari, Leena Uotila and Riitta Velin

Oral examination: Kirsi Ihanus (dentist), Ulla Palotie (dentist) and Sari Salminen (field nurse)

Courage-nurses: Mira Kokko, Hannele Lehtonen and *Anne Lunkka

Tampere

Field nurses: Elina Poutanen (head nurse), Jaana Haapanen, Anneli Hellstrom, Johannes Hietala, Päivi Hyvärinen, Jonna Järvitalo, Leena Kangas-Viri, Krista Karjalainen, Marjaana Leinonen, Hanna-Mari Lyömilä, *Liisa Olkkonen, Anneli Peltonen, Minna Putkisaari, Anne Silen and Marja Sipola

Courage-nurses: Marja Ala-Lahti, Anna Kupila and Eeva-Maria Niemelä

Kuopio

Field nurses: Ulla Tyyni (head nurse), Minna Haapalainen, Liisa Heiskanen, *Helena Holopainen, Anna-Kaarina Ikonen, Raija-Leena Juntunen, Laura Karvonen, Johanna Kemppainen, Päivi Kettunen, Anna Sofia Kuoppa, Anri Leppänen, Virpi Putkonen, Marita Ronkainen, Päivi Viiliäinen and Leila Väänänen

Courage-nurses: Titta Luomanmäki-Rantala, Carita Röpelinen-Pitkänen and Santtu-Pekka Turunen

Oulu

Field nurses: Liisa Moilanen (head nurse), Arja Halmetoja, Riitta Höyhtyä, Lea Jämsä, *Kirsi Kariniemi, Marika Kiviniemi, Katrianna Kyllönen, Minna Mattlar, Jaana Pakanen, Seija Riekki, Anu Räisänen, Anne Sarajärvi, Marjut Savuoja, Anna-Maria Silenius, Marja Suorsa, Arja Tapio and Riitta Tervo

Oral examination: Jenni Myöhänen (dentist), Petra Timonen (dentist) and Arja Alaviuhkola (field nurse)

Courage-nurses: Saija Glader, Tiina Hanhikorpi and Maija Norman

Turku

Field nurses: Marjatta Härkäs (head nurse), Aniitta Entonen, Tuija Heinonen, Viljami Hätönen, Mia Horkko, Ari Järvensivu, *Ritva Kivi, Pihla Laakso, Ulla Paassilta, Arja Parkkonen, Henna-Riikka Ratola, Matias Raunio, Paula Ronkainen, Minna Suominen, Eija Tervola and Kaija Varjonen

Courage-nurses: Salla Miesvirta and Maarit Seitz

* participated also as a field nurse in the complementary data collection

APPENDIX 3. Research locations, periods and sample sizes by fieldwork teams

Research location (sample size)	Sample municipalities included in the sample size	Period
Fieldwork team 1		
Helsinki (2066)	Helsinki, Vantaa, Espoo, Nurmijärvi, Kirkkonummi, Kauniainen	10.08.2011-08.12.2011
Porvoo (103)	Porvoo, Askola, Pornainen	09.12.2011-15.12.2011
Loviisa (77)	Loviisa, Lapinjärvi	16.12.2011-21.12.2011
Fieldwork team 2		
Tampere (594)	Tampere, Nokia, Ylöjärvi, Kangasala, Lempäälä, Pirkkala, Orivesi, Sastamala (Vammala), Hämeenkyrö, Ikaalinen, Parkano, Ruovesi, Juupajoki	10.08.2011-14.09.2011
Seinäjoki (224)	Seinäjoki, Ilmajoki, Kauhajoki, Kurikka, Alavus, Jalasjärvi, Kihniö	15.09.2011-29.09.2011
Lapua (97)	Lapua, Kauhava, Alajärvi, Lappajärvi, Vimpeli	30.09.2011-06.10.2011
Keuruu (78)	Keuruu, Multia, Virrat, Mänttä-Vilppula	07.10.2011-12.10.2011
Valkeakoski (108)	Valkeakoski, Akaa (Toijala), Pälkäne, Urjala, Vesilahti	13.10.2011-20.10.2011
Hämeenlinna (119)	Hämeenlinna, Janakkala, Hattula	21.10.2011-28.10.2011
Heinola (94)	Heinola, Hartola, Sysmä	31.10.2011-04.11.2011
Lahti (241)	Lahti, Hollola, Nastola, Asikkala, Padasjoki, Hämeenkoski, Kärkölä	07.11.2011-21.11.2011
Orimattila (91)	Orimattila, Pukkila, Myrskylä	22.11.2011-28.11.2011
Riihimäki (112)	Riihimäki, Hausjärvi, Loppi	29.11.2011-08.12.2011
Hyvinkää (99)	Hyvinkää, Mäntsälä	09.12.2011-15.12.2011
Fieldwork team 3		
Turku (614)	Turku, Kaarina, Naantali, Masku, Nousiainen, Raisio, Rusko, Lieto, Paimio, Sauvo, Mynämäki, Maarianhamina, Eckerö	10.08.2011-15.09.2011
Länsi-Turunmaa (Parainen) (51)	Länsi-Turunmaa (Parainen)	16.09.2011-20.09.2011
Uusikaupunki (83)	Uusikaupunki, Vehmaa, Taivassalo, Kustavi, Laitila	21.09.2011-27.09.2011
Rauma (51)	Rauma, Eurajoki	28.09.2011-30.09.2011
Harjavalta (97)	Harjavalta, Nakkila, Eura, Huittinen, Kokemäki	03.10.2011-07.10.2011

Kristiinankaupunki (86)	Kristiinankaupunki, Kaskinen, Karvia, Isojoki, Karijoki, Teuva	10.10.2011-14.10.2011
Pori (180)	Pori, Ulvila, Merikarvia, Kankaanpää	17.10.2011-27.10.2011
Loimaa (53)	Loimaa, Pöytyä, Oripää, Aura, Ypäjä	28.10.2011-01.11.2011
Somero (47)	Somero	02.11.2011-04.11.2011
Forssa (91)	Forssa, Tammela, Jokioinen, Humppila	07.11.2011-11.11.2011
Karkkila (100)	Karkkila, Vihti	14.11.2011-18.11.2011
Salo (159)	Salo, Kemiönsaari	21.11.2011-30.11.2011
Lohja (108)	Lohja, Nummi-Pusula, Raasepori, Siuntio, Inkoo, Hanko, Karjalohja	01.12.2011-12.12.2011
Tuusula (117)	Tuusula, Järvenpää, Kerava, Sipoo	13.12.2011-20.12.2011
Fieldwork team 4		
Kuopio (197)	Kuopio, Karttula, Leppävirta, Varkaus, Suonenjoki, Tervo, Kaavi, Tuusniemi	10.08.2011-22.08.2011
Siilinjärvi (83)	Siilinjärvi, Maaninka, Juankoski, Nilsiä	23.08.2011-29.08.2011
Juuka (69)	Juuka	30.08.2011-02.09.2011
Polvijärvi (57)	Polvijärvi, Outokumpu	05.09.2011-07.09.2011
Joensuu (181)	Joensuu, Kontiolahti, Liperi, Ilomantsi, Kitee, Rääkkylä, Heinävesi	08.09.2011-19.09.2011
Lieksa (74)	Lieksa	20.09.2011-23.09.2011
Savonlinna (156)	Savonlinna, Kerimäki, Enonkoski, Punkaharju, Rantasalmi, Sulkava, Parikkala	26.09.2011-05.10.2011
Mikkeli (102)	Mikkeli, Ristiina, Hirvensalmi, Kangasniemi, Pertunmaa, Mäntyharju, Pieksämäki, Puumala	06.10.2011-12.10.2011
Jyväskylä (248)	Jyväskylä, Muurame, Laukaa, Petäjävesi, Viitasaari, Luhanka, Äänekoski, Joutsa, Karstula, Kyyjärvi, Toivakka, Uurainen	13.10.2011-27.10.2011
Jämsä (76)	Jämsä, Kuhmoinen	28.10.2011-2.11.2011
Imatra (139)	Imatra, Ruokolahti	03.11.2011-11.11.2011
Luumäki (68)	Luumäki	14.11.2011-17.11.2011
Lappeenranta (119)	Lappeenranta, Taipalsaari, Savitaipale, Lemi	18.11.2011-25.11.2011
Kouvola (94)	Kouvola, Iitti	28.11.2011-02.12.2011
Hamina (83)	Hamina, Virolahti	07.12.2011-13.12.2011
Kotka (92)	Kotka, Pyhtää	14.12.2011-20.12.2011

Fieldwork team 5		
Oulu (465)	Oulu, Haukipudas, Kiiminki, Ii, Kempele, Liminka, Muhos, Oulunsalo, Tyrnävä, Lumijoki, Siikalatva, Utajärvi, Yli-Ii	10.08.2011-06.09.2011
Raahe (68)	Raahe, Siikajoki, Pyhäjoki	07.09.2011-12.09.2011
Kokkola (80)	Kokkola, Kaustinen, Toholampi, Luoto, Kruunupyy	13.09.2011-16.09.2011
Uusikaarlepyy (183)	Uusikaarlepyy, Pietarsaari, Pedersöre	19.09.2011-29.09.2011
Vaasa (107)	Vaasa, Mustasaari, Isokyrö, Maalahti, Vähäkyrö, Vöyri	03.10.2011-10.10.2011
Ylivieska (66)	Ylivieska, Alavieska, Kalajoki, Oulainen, Sievi	11.10.2011-14.10.2011
Nivala (55)	Nivala, Haapajärvi, Reisjärvi	17.10.2011-19.10.2011
Pyhäjävi (50)	Pyhäjärvi	20.10.2011-24.10.2011
Iisalmi (149)	Iisalmi, Lapinlahti, Vieremä, Sonkajärvi, Rautavaara	25.10.2011-02.11.2011
Kajaani (67)	Kajaani, Puolanka, Ristijärvi, Kuhmo, Vaala	03.11.2011-08.11.2011
Kuusamo (56)	Kuusamo, Posio	09.11.2011-11.11.2011
Taivalkoski (59)	Taivalkoski, Pudasjärvi	14.11.2011-17.11.2011
Utsjoki (58)	Utsjoki, Inari	21.11.2011-23.11.2011
Sodankylä (57)	Sodankylä, Kittilä, Pelkosenniemi, Kolari	24.11.2011-26.11.2011
Rovaniemi (78)	Rovaniemi, Pello, Ranua, Kemijärvi	29.11.2011-02.12.2011
Kemi (94)	Kemi, Simo, Tornio, Keminmaa, Tervola	07.12.2011-13.12.2011

APPENDIX 4. Contents of interviews, short questionnaire and questionnaire for young adults (following the structure of the long interview)

	Long interview	Short interview	Interview for young adults	Phone interview	Short question-naire	Question-naire for young adults
A. Background information						
AA. (Mother tongue,) marital status and relationship	AA00-06	AA00-01	AA00-06	AA01	AA01	AA01
AB. Household	AB01	AB01	AB01	AB01	AB01	AB01
AC. Education	AC01-03	AC01-02	AC01N-03	AC01-02	AC01-02	AC01N-02N
AD. Main activity, occupation	AD01-08	AD01-05	AD01-08	AD01-02, AD04	AD01-02, AD04	AD01
AE. Present/previous occupation (main job)	AE01-09	AE01-06	AE01-09	AE01	AE01	AE01
AF. Working hours and income (main job)	AF01-05	-	AF01-05	-	-	-
AF. Unemployment	AH01-05	AH01-03	AH01-05	AH01, AH03	AH01, AH03	AH01-03
AI. Information about your spouse	AI04-05	AI04-05	AI04	AI04-05	-	-
B. Health and illnesses						
BA. Perceived health and chronic illnesses	BA01-03	BA01-02	BA01-03	BA01-02	BA01-02	BA01-02
Respiratory diseases	BA04-07	BA04-06	BA04	BA04-06	BA04-06	BA04
Heart diseases	BA08-12	BA08-10	-	BA08-10	BA08-10	-
Other diseases in the vascular system	BA13-16	BA13-14	BA13	BA13-14	BA13-14	-
Defects, diseases and injuries of the joints, extremities and the back	BA18-24	BA18-24	BA18-24	BA18-19, BA21,BA24	BA18-19, BA21, BA24	BA24
Mental health problems	BA25	BA25	BA25	BA25	BA25	BA25
Vision and hearing defects	BA36-40	BA36-40	BA39-40	BA36-38, BA40	BA36-38, BA40	-

	Long interview	Short interview	Interview for young adults	Phone interview	Short question-naire	Question-naire for young adults
Other diseases diagnosed by a doctor	BA26-46	BA26-46	BA26-46	BA26, BA32, BA35, BA44-46	BA26, A32, BA35,BA44, BA46	BA26-27, BA46
BB. Treatment of illnesses	BB01	BB01	BB01	BB01	BB01	-
Hospital care	BB10	BB10	-	BB10	BB10	-
Surgical operations	BB12	-	-	-	-	-
BC. Questions for men	BC02	BC02	BC02	-	BC02	BC02
Infertility	BC03-04	-	BC03-04	-	BC03-04	-
Contraception	BC05	-	BC05	-	-	BC05
BD. Questions for women	BD00	-	BD00	-	-	-
Menstruation	BD02-06	-	BD02-05	-	-	-
Pregnancies and deliveries	BD07-22	BD12, 22	BD12-22	-	BD08, 19-21	BD08, 19-22
Childlessness and infertility treatments	BD23-25	-	BD23-25	-	BD23-24	-
Contraception	BD26, BD29	-	BD26, 29	-	-	BD26
Hormone replacement therapies	BD32-34	-	-	-	BD32	-
C. Questions concerning (your) parents and siblings						
A. Parents	-	-	CA01-11	-	-	CB01_2_T 11
CB. Living conditions in childhood	-	-	CB01-11	-	-	CB05-06, 10-11
D. Health services						
DA. Availability and accessibility	DA00-07	DA00-04	DA00-07	-	-	-
DB. Ambulatory visits due to illnesses or symptoms	DB01-11	DB01, 08	DB01-11	DB01	DB01, DB03, DB08	DB0,DB03, DB08

	Long interview	Short interview	Interview for young adults	Phone interview	Short question-naire	Question-naire for young adults
DC. Mental health services	DC01-12	DC01-02, 05	DC01-12	-	-	DC00AN, DC10N
DD. Health examinations and preventive health services	DD01-08	DD04-08	DD01-04	-	-	DD01
DE. Physiotherapy	DE01-03	DE01-02	DE01-03	-	-	-
Other treatments and therapies	DE03-04	-	DE03-04	-	-	-
DF. Medicines	DF01-09	DF01-09	DF01-09	DF02-03	DF02, DF03	DF02-04
E. Oral health						
EA. Oral health status	EA01-05	EA01-05	EA01-02	-	-	-
EB. Oral self-care	EB01-03	EB01	EB01-03	-	-	-
EC. Use of services	EC01, EC05-06	EC01, 06	EC01, EC05-06	EC06	EC06	EC06
ED. Dental care customer	ED01-07	ED04	ED01-07	ED04	ED04	ED04
F. Living habits						
FA. Eating habits	FA02-12	FA02-11	FA05-12	-	FA09	-
FB. Tobacco	FB01-08	FB01-07	FB01-08	FB05	FB05	FB05
Height and weight (see Questionnaire 1, T4002)	-	-	-	Kys1_K17, Kys1_K17b Kys1_K28, T11_Kys1_ K28	Kys1_K17, Kys1_K17b Kys1_K28, T11_Kys1_K2 8	Kys1_K17, Kys1_K17b Kys1_K28, T11_Kys1_ K28
Exercise (see Questionnaire 1, T4002)	-	-	-	K22	K1_K21-22	-
Use of alcohol (see Questionnaire 1, T4002)	-	-	-	K28	K1_K28	-
G. Living environment						
GB. Residential environment	GB01, GB06	GB01, 06	GB04	-	-	-
Hindrances in your residential environment	GB05-10	GB05-10	-	-	-	-
GC. Neighbourhood services	GC01-07	-	-	-	-	

	Long interview	Short interview	Interview for young adults	Phone interview	Short question-naire	Question-naire for young adults
H. Functional capacity						
HA. Activities of daily living (ADL and IADL)	HA01-02	HA01-02	HA01	HA01-02	HA01-02	-
HB. Mobility	HB01-14	HB01-11	HB01-08	HB02, HB6-7, HB9-10	HB06-07, HB09-10, HB02	HB01, HB03
HC. Sensory functions	HC01-06	HC01-06	HC01-04	HC01, HC04, HC06	HC01, HC04	HC01, HC04
HD. Need and use of assistance and help	HD01-10	HD01-10	HD01-10	HD01-03, HD05-06, HD08-10	HD01-03, HD05-06, HD08-10	-
HG. Aids	HG01-09	HG01-09	HG01-09	HG01-03, HG05, HG07	HG01-03, HG05, HG07	HG01
HF. Cognitive capacity	HF01-14	HF01-14	HF09-10	HF09-10	HF09-10	HF09_A, HF09_B, HF10
I. Work and working ability						
IA. Working conditions	IA04	-	IA04	-	-	-
IB. Work ability	IB01-15	IB01-12	IB01-15	IB01, IB04, IB11-12	IB01, IB04, IB11-12	IB01, IB04
IC. Skills	IC01-04	-	IC01-04	-	-	-
ID. Retirement attitudes	ID01-02	-	-	-	-	-
IE: Working history	IE01-06	-	IE01-06	-	-	-
J. Rehabilitation						
JA. Use of services	JA01, JA03, JA07	JA01	JA01, JA03, JA07	-	-	-
JB. Need for rehabilitation	JB01, JB03, JB06	JB01-06	JB01, JB03, JB06	-	-	-

	Long interview	Short interview	Interview for young adults	Phone interview	Short question-naire	Question-naire for young adults
K. Interviewer's assessments						
KA. Health examination	KA01	KA01	KA01	KA01	-	-
KB. Interviewer's assessments on the factors affecting the interview	KB03-06	KB03-06	KB03-06	KB04-06	-	-

APPENDIX 5. Contents of the main questionnaires

QUESTIONNAIRE 1

	Question number in T4002	Question number in T4002_2 (young adults)	Instrument
Childhood	-	63	
Eating or drinking sweets or sweetened drinks	39	38	
Eating problems	-	39-43	SCOFF
Exercise	21-25	22-26	
Functional capacity	1-3	1-2	EuroQol
Health promotion	40	49	
Household income and expenditure	53	62	
Loneliness	52	61	
Neighbourhood safety	5-6	4-5	
Percenption and strain of work and studies	-	53	Maslach (MBI)
Perception and strain of work	44-46	-	Maslach (MBI)
Psychological well-being	7-18	6-17	GHQ
Quality of life	41-43	50-52	WHOQOL
Self-harm	-	18-19	
Sexuality	-	44-48	
Social environment	4	3	
Time management and hobbies	26	27	
Treatment of drinking problems	37-38	-	
Use of alcohol	27-36	28-37	AUDIT
Weight and height	19-20	20-21	
Vitality	49-51	58-60	
Working conditions	47-48	54-57	

QUESTIONNAIRE 2

	Question number in T4003	Question number in T4003_2 (young adults)	Instrument
Allergic symptoms	14-19	1-6	
Back pain	20-30	7-17	
Chest pain	5-13	-	
Falling	51-53	-	
Joint pain	45-46	32-33	
Mood and feelings	54	36	BDI-13 (BECK)
Musculoskeletal symptoms			
Neck pain	31-39	18-26	
Orientation to life	55	37	Antonovsky SOC-13
Overall hindrance	47-48	34-35	
Shortness of breath	1-4	-	
Shoulder pain	40-44	27-31	
Walking difficulties	49-50	-	

QUESTIONNAIRE 3

	Question number in T4003	Question number in T4003_2 (young adults)	Instrument
Emotions and feelings	14	9	TAS-20
Financing of care and health services	15-18	-	
Health related quality of life	11	7	15D
Oral health and quality of life	12	-	OHIP
Relationships	13	8	
Seasonal variation	8-9	4-5	
Sleep	1-7	1-3	
Symptoms	10	6	SCL-25

QUESTIONNAIRE 4 (T4045)

	Question number	Origin of the question(s)
Health status	1	Questionnaire T4002, no. 2
	2	Questionnaire T4003, no. 14
	3	Questionnaire T4003, no. 24
	4	Questionnaire T4003, no. 33
	5	Questionnaire T4003, no. 42
	6-11	Questionnaire T4003, no. 45-50
Living habits	12	Questionnaire T4002, no. 21
	13	Questionnaire T4002, no. 28
Quality of life	14	Questionnaire T4002, no. 41
	15-16	Questionnaire T4002, no. 52-53

SHORT QUESTIONNAIRE (T4095)

	Question number	Origin of the question(s)
Aids	54	Interview T4001: HG01-03, HG05, HG07
Cardiovascular diseases	6-10	Interview T4001: BA08-10, BA13-14
Defects and diseases of the joints and back	11-14	Interview T4001: BA18-19, BA21
Dental care	31-32	Interview T4001: EC06, ED04
Education	67-68	AC01-02
Family	62-66	AA01, AB01, BD22-24
Functional capacity	37-39	Interview T4001: HA01-02, HB06-07, HB09-10, HB02
(sight)	40	Interview T4001: HC01
(hearing)	41	Interview T4001: HC04
(cognitive capacity)	42-44	Interview T4001: HF09-10
Health and illnesses	1-2	Interview T4001: BA01-02
Height and weight	55-56	Questionnaire T4002: no. 19-20
Injuries	15	Interview T4001: BA24
Living habits	57-61	
(smoking)	57	FB05
(exercise)	58-59	K1_K21-22
(vegetables)	60	FA09
(alcohol)	61	K1_K28
Main activity and occupation	69-74	AD01-02, AD04, AE01, AH01, AH03
Medicines	30	Interview T4001: DF02, DF03
Mental health problems	16	Interview T4001: BA25
Need and use of assistance and help	45-53	Interview T4001: HD01-03, HD05-06, HD08-10
Other diseases diagnosed by a doctor	21-25	Interview T4001: BA26, BA32, BA35, BA44, BA46
Participation in a health survey	75-76	-
Questions for women	77-80	BD08, BD19-21, BD32
Respiratory diseases	3-5	Interview T4001: BA04-06
Treatment of illnesses	26-27	Interview T4001: BB01
Use of health care services	28-29	Interview T4001: BB10, DB01, DB03, DB08
Vision and hearing defects	17-20	Interview T4001: BA36-38, BA40
Work capacity	33-36	Interview T4001: IB01, IB04, IB11-12

QUESTIONNAIRE FOR YOUNG ADULTS (T4140)

	Question number	Origin of the question(s)
Ambulatory visits due to illnesses or symptoms	23-30	Interview T4142: DB01, DB03, DB08, DD01, EC06, ED04, DC12
Childhood	5-9	Interview T4142: CB05-06, CB10-11; Questionnaire T4002_2: no. 63
Contraception	68-69	Interview T4142: BC05/BD261, BC05_1/BD261_1b; Questionnaire T4002_2: no. 48
Education	10-11	Interview T4142: T4142: AC01-02
Exercise	70-72	Interview T4142: HB01, HB03; Questionnaire T4002_2: no. 23
Family	1-4	Interview T4142: AA01, AB01, BC02/BD22
Health	15-22	Interview T4142: BA01-02, BA04; Questionnaire T4003_2: no. 2, 4, Interview T4142: BA26; BA24; BA25, Ba25a; BB01A, BB01C
Main activity and occupation	73-75	Interview T4142: AD01, AE06a, AE01
Medicines	48-50	Interview T4142: DF02-04
Memory and learning	59-61	Interview T4142: HF09_A, HF09_B, HF10
Psychological well-being	31-45	Questionnaire T4002_2: no. 6-17 /GHQ), 61, 18-19
Quality of life	12-14	Questionnaire T4002_2: no. 50-52
Questions for women	79-81	Interview T4142: BD08d, BD21, BD19&20 (combined)
Sleep and sleeping	46-47	Questionnaire T4005_2: 1-2
Smoking and alcohol	64-67	Interview: T4142: FB05; Questionnaire 4002_2; no 29-31
Symptoms	51-53	Questionnaire T4003_2: 9, 18, 27, 11, 20, 29 (headache question not found); 53: not found
Unemployment	76-78	Interview T4142: AH01-03
Weight and height	62-63	Questionnaire T4002_2: no. 20-21
Vision and hearing	56-58	Interview T4142: HG01, HC01, HC04
Work capacity	54-55	Interview T4142: IB01, IB04

APPENDIX 6. Contents of the COURAGE in Europe sub-study

Interview and examinations at home	
Composition and socioeconomic position of the household	Q0401
Sociodemographic characteristics	Q1008-1024
Work history and benefits	Q1501-1516
Health state descriptions	
Self-rated health	Q2000-2001
Pain and discomfort	Q2007-2009
Mobility	Q2003-2046
Self-care	Q2037-2006
Cognition	Q2035-2011
Interpersonal activities	Q2032-2039
Domestic life and work	Q2032-2039
Sleep and energy	Q2016-2017
Affect	Q2018-2048
Vision	Q2020-2024
Hearing	Q2050-2049
Measurements and performance tests	
Blood pressure and heart rate	Q2501-2503
Distance and near vision	Q2514-2517
Digit span: digits forward and backward	Q2534-2535
Grip strength	Q2518-2522
Risk factors and preventive health behaviours	
Smoking	Q3001-3006
Alcohol	Q3007-3011
Nutrition	Q3012-3015
Physical activity	Q3016-3031
Chronic conditions and health services coverage	
Joint disorders and back pain	Q4001-4009
Stroke	Q4010-4013
Angina	Q4014-4021
Diabetes	Q4022-4024
Chronic lung disease	Q4025-4032
Asthma	Q4033-4039
Depression	Q4040-4059
Hypertension	Q4060-4061

Cataract	Q4062-4065
Oral health	Q4066-4068
Injuries	Q4069-4077
Cervical cancer and breast cancer screening	Q4078-4080
Health care utilization	
Latest contact with health care	Q5001-5004
Inpatient hospital care	Q5005-5008
Outpatient care and care at home	Q5026-5029
Health expenditure	Q0804-0810
Social cohesion and social network	
General questions	Q6010-6117
Communication	Q6200-6230
Support	Q6300-6353
Trust	Q6400-6440
Safety	Q6500-6520
Participation	Q6600-6640
Subjective well-being and quality of life	
Quality of life	Q7001-7010
Stress	Q7008-7010
Day reconstruction	Q7011-7533
Built environment	
General information	Q8101
Neighbourhood environment	Q8201-8213
Open-to-public buildings, places and facilities	Q8301-8307
Living place / home	Q8401-8410
Overall questions	Q8510-8530
Interviewer assessment	Q9001-9014

Built environment outdoor checklist, filled in by the fieldwork personnel	
Streetscape	Q11101-11602
Walkways	Q12110-12610
Bikeways	Q13100-13110
Street crossing/intersections	Q14100-14110
Parking facilities	Q15101-15107
Public facilities and features of the street	Q16101-16106
Land-use visible along the street/road	Q17101-17106
Site decay/urban blight	Q18101-18108
Street activity	Q19101-19107